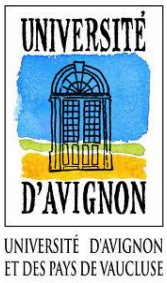
Euclidean Jordan algebras and variational problems under conic constraints

David Sossa



ACADÉMIE D'AIX-MARSEILLE
UNIVERSITÉ D'AVIGNON ET DES PAYS DE VAUCLUSE

THÉSE DE DOCTORAT

présentée á l'Université d'Avignon et des Pays de Vaucluse
pour l'obtention du grade de Docteur

SPECIALITE : Mathématiques

Algèbres de Jordan euclidiennes et problèmes variationels avec contraintes coniques

par
David SOSSA

soutenue le 4 septembre 2014 devant un jury composé de

A. DANIILIDIS	Université du Chili	Rapporteur
P. GAJARDO	Université Federico Santa María	Examinateur
S. GOWDA	Université du Maryland	Rapporteur
D.T. LUC	Université d'Avignon	Examinateur
H. RAMÍREZ	Université du Chili	Directeur
A. SEEGER	Université d'Avignon	Directeur
M. VOLLE	Université d'Avignon	Examinateur

Résumé

Cette thèse concerne quatre thèmes apparemment différents, mais en fait étroitement liés: des problèmes variationnels sur les algèbres de Jordan euclidiennes, des problèmes de complémentarité sur l'espace des matrices symétriques, l'analyse angulaire entre deux cônes convexes fermés et analyse du chemin central en programmation conique symétrique.

Dans la première partie de ce travail, le concept de "commutation au sens opérationnel" dans les algèbres de Jordan euclidiennes est étudié en fournissant un principe de commutation pour les problèmes variationnels avec des données spectrales.

Dans la deuxième partie, nous abordons l'analyse et la résolution numérique d'une large classe de problèmes de complémentarité sur l'espace des matrices symétriques. Les conditions de complémentarité sont exprimées en termes de l'ordre de Loewner ou, plus généralement, en termes d'un cône du type Loewnerien.

La troisième partie de ce travail est une tentative de construction d'une théorie générale des angles critiques pour une paire de cônes convexes fermés. L'analyse angulaire pour une paire de cônes spécialement structurés est également considérée. Par-exemple, nous travaillons avec des sous-espaces linéaires, des cônes polyédriques, des cônes de révolution, des cônes "topheavy" et des cônes de matrices.

La dernière partie de ce travail étudie la convergence et le comportement asymptotique du chemin central en programmation conique symétrique. Ceci est fait en utilisant des techniques propres aux algèbres de Jordan.

Abstract

This thesis deals with four different but interrelated topics: variational problems on Euclidean Jordan algebras, complementarity problems on the space of symmetric matrices, angular analysis between two closed convex cones and the central path for symmetric cone linear programming.

In the first part of this work we study the concept of "operator commutation" in Euclidean Jordan algebras by providing a commutation principle for variational problems involving spectral data.

Our main concern of the second part is the analysis and numerical resolution of a broad class of complementarity problems on spaces of symmetric matrices. The complementarity conditions are expressed in terms of the Loewner ordering or, more generally, with respect to a dual pair of Loewnerian cones.

The third part of this work is an attempt to build a general theory of critical angles for a pair of closed convex cones. The angular analysis for a pair of specially structured cones is also covered. For instance, we work with linear subspaces, polyhedral cones, revolution cones, topheavy cones and cones of matrices.

The last part of this work focuses on the convergence and the limiting behavior of the central path in symmetric cone linear programming. This is done by using Jordan-algebra techniques.

Resumen

En esta tesis doctoral se abordan cuatro tópicos diferentes pero mutuamente relacionados: Problemas variacionales sobre álgebras de Jordan Euclideanos, problemas de complementariedad sobre espacios de matrices simétricas, análisis angular entre dos conos convexos y cerrados, y el camino central en programación cónica simétrica.

La primera parte de este trabajo corresponde al estudio del concepto de "operator commutation" en álgebras de Jordan Euclideanos por medio del establecimiento de un principio de conmutación para problemas variacionales los cuales poseen datos espectrales.

El principal enfoque de la segunda parte es el análisis y resolución numérica de una amplia clase de problemas de complementariedad formuladas en espacios de matrices simétricas. Las condiciones de complementariedad son expresadas en términos de la ordenación de Loewner o, mas general, con respecto a un par dual de conos Loewnerianos.

En la tercera parte presentamos una construcción de la teoría general de ángulos críticos para pares de conos convexos y cerrados. El análisis angular de pares de conos con estructuras especiales es también abordada. Por ejemplo, en nuestro estudio incluimos: subespacios lineales, conos poliedrales, conos de revolución, conos topheavy y conos de matrices.

La última parte de este trabajo está dedicada al estudio de la convergencia del camino central y del comportamiento de su punto límite en programación cónica simétrica. Esto es hecho por medio del uso de herramientas de álgebras de Jordan.

A mis padres Yola y Victoriano,
a mi esposa Liz Fabiola
y a mi hija Flor Daliz.
Con mucho Amor.

Acknowledgments

I thank God for blessing me much more than I deserve and for his wonderful promise expressed in John 3:16 "*For God so loved the world, that he gave his only begotten Son, that whosoever believeth in him should not perish, but have everlasting life.*"

I would like to thank my two advisors, Prof. Héctor Ramírez and Prof. Alberto Seeger. Their patience, encouragement, and immense knowledge were key motivations throughout my PhD. It is a real honor for me to have worked with them. Special thanks for the friendship and the hospitality received during my stay in Santiago and Avignon.

I would like to thank Prof. Aris Daniilidis and Prof. Seetharama Gowda for review the earlier version of this thesis and for several suggestions that improved the presentation. I am also grateful to my oral committee members: Prof. Pedro Gajardo, Prof. Dinh The Luc and Prof. Michel Volle for their time, interest, helpful comments and insightful questions.

I would also like to express my gratitude to all my friends I made during this period of time at the DIM-UCH and at the LMA-UAPV. Special thanks to the couples Paola and César, Ying and Huyuan. Thank you for all of your support and sincere friendship. Special thanks to Caroline for her help and warm hospitality offered during my stay in Avignon.

I gratefully acknowledge the funding sources that made my Ph.D. work possible. My studies in Chile were funded by CONICYT (Comisión Nacional de Investigación Científica y Tecnológica de Chile) and during my stay in France I was funded by the grants for short research stays from the Universidad de Chile, from the Center for Mathematical Modeling (CMM) and from the Embassy of France in Chile.

Last but not least, I thank to my family. Thanks to my beloved wife Liz Fabiola and my beloved daughter Flor Daliz. Thank you for being there every step of the way with me. You were the key motivation for this achievement. Thanks for filling my life with happiness. Thanks to my dear parents Yola and Victoriano for their support, prayers and love. Thanks to my sisters Ana, Isabel, Marcela, Verónica and my brothers René and Efraín for be with me in all time.

Contents

Introduction

This doctoral thesis deals with variational problems over Euclidean Jordan algebras and variational problems involving conic constrains. It is composed by five chapters which are based on the papers [12, 15] and the manuscripts [16, 17, 13], respectively. In order to facilitate the reading of this work, each chapter is presented in a self-contained way.

In the first chapter we give some contributions to the theory of Euclidean Jordan algebras by establishing a commutation principle for variational problems formulated in this context.

The second chapter deals with the analysis and numerical resolution of complementarity problems with respect to Loewnerian cones.

The concept of critical angles between two closed convex cones is developed in the third and fourth chapter. The general theory is presented in the third chapter and some particular structures are studied in the fourth chapter.

In the last chapter we present some results concerning to the convergence and the limiting behavior of the central path in symmetric cone linear programming.

A brief description of the main problems that we have considered and the main results that we have obtained is presented in the next paragraphs.

Commutation principle for variational problems on Euclidean Jordan algebras

In this chapter we work in the context of Euclidean Jordan algebras (EJA). A preliminary material concerning to the theory of EJA is given on Section 1.2 of Chapter 1 and Section 5.2.1 of Chapter 5. For more details we refer to the books [3, 10].

Let $(\mathbb{V}, \circ, \langle \cdot, \cdot \rangle)$ be a EJA with rank r. One says that two elements $a, b \in \mathbb{V}$ *operator commute* if they satisfy

$$a \circ (b \circ z) = b \circ (a \circ z) \quad \text{for all } z \in \mathbb{V}.$$

The operator commutative property is highly useful to deduce some important results in EJA (e.g. [1, 11]). A set $\Omega \subseteq \mathbb{V}$ is called *spectral set* if there exists a permutation invariant set

$Q \subseteq \mathbb{R}^r$ such that

$$\Omega = \{x \in \mathbb{V} : \lambda(x) \in Q\},$$

where $\lambda(x) \in \mathbb{R}^r$ is the vector of the eigenvalues of x arranged in nondecreasing order. A function $\Phi : \mathbb{V} \to \mathbb{R}$ is called *spectral function* if there exist a permutation invariant function $g : \mathbb{R}^r \to \mathbb{R}$ such that

$$\Phi(x) = g(\lambda(x)).$$

The main contribution presented in this chapter is the establishment of a commutation principle for variational problems involving spectral data. It reads as follows:

Theorem. *Let $(\mathbb{V}, \circ, \langle\cdot,\cdot\rangle)$ be a Euclidean Jordan algebra and let $a, b \in \mathbb{V}$. Suppose that $\Omega \subseteq \mathbb{V}$ is a spectral set and that $\Phi : \mathbb{V} \to \mathbb{R}$ is a spectral function. Under these assumptions, if b is a local minimum or a local maximum of*

$$x \in \Omega \mapsto F(x) = \langle a, x\rangle + \Phi(x),$$

then a and b operator commute.

This result was inspired by the commutation principle established by Iusem and Seeger [7, Lemma 4] for the particular case of the space of symmetric matrices.

Complementarity problems with respect to Loewnerian cones

This chapter is devoted to the analysis and numerical resolution of a class of nonlinear complementarity problems formulated in $\mathbb{S}^n$, the space of symmetric matrices of order n. The complementarity problem studied is of the form

$$\begin{cases} \Phi(X, Y, \lambda) = \mathbf{0} \\ \mathcal{K} \ni X \perp Y \in \mathcal{K}^*, \end{cases} \tag{1}$$

where λ stands for an unknown parameter vector in some Euclidean space, Φ is a continuously differentiable function and $\mathcal{K}$ is a Loewnerian cone, i.e., it is the image of the positive semidefinite cone $\mathbb{S}^n_+$ under some linear endomorphism on $\mathbb{S}^n$.

We show that, under some assumptions over the dimension of the involved spaces, the problem (1) can be equivalently formulated as a square system of nonlinear equations by using complementarity functions for the cone $\mathbb{S}^n_+$. In this work we make use of the *Fischer-Burmeister* complementarity function and the *Minimum* complementarity function. Hence, for solving (1) we apply the Semismooth Newton Method to the equivalent formulation.

The performance of these methods is studied by considering the particular case when $\mathcal{K} = \mathbb{S}^n_+$. The problem tested in our numerical experiments is that of finding a Loewner-eigenvalue of a linear map generated by means of a random mechanism.

As a theoretical result of this chapter we study some properties of Loewnerian cones and we provide a list of examples of Loewnerian cones and counter-examples.

We end this chapter by showing that the problem of finding the nearest Euclidean distance matrix can be formulated as a complementarity problem of the form (1). We also report the performance of our numerical methods for solving this particular problem.

Critical angles between two convex cones I. General theory

Given two nontrivial closed convex cones P and Q in a Euclidean space $\mathbb{X}$, we define the *maximal angle* between P and Q as the number

$$\Theta(P,Q) \ := \max_{u\in P\cap S_{\mathbb{X}},\, v\in Q\cap S_{\mathbb{X}}} \arccos\langle u,v\rangle, \tag{2}$$

where $S_{\mathbb{X}}$ stands for the unit sphere of $\mathbb{X}$. A *critical pair* of (P,Q) is a pair $(u,v)\in\mathbb{X}^2$ satisfying the following necessary optimality conditions for the nonconvex optimization problem (2):

$$u\in P\cap S_{\mathbb{X}},\ v\in Q\cap S_{\mathbb{X}},\ v-\langle u,v\rangle u\in P^*,\ u-\langle u,v\rangle v\in Q^*.$$

The corresponding angle $\arccos\langle u,v\rangle$ is called a *critical angle* of (P,Q). The critical pair that achieves the maximal angle is called *antipodal pair*.

In this chapter we attempt to build a general theory of critical angles for a pair of closed convex cones. Most of the results presented in this chapter is inspired by the recent theory of critical angles for a closed convex cone, as developed in [8].

A boundary principle is established which help us to understand where the critical pairs are localized. Further characterization of criticality and antipodality are also provided. As a motivation of the study of the concept of maximal angle we show how the pointedness and reproducibility properties of a pair of closed convex cones are related with the maximal angle. Some continuity issues are also covered.

We also discuss the particular case of critical angles in a pair of linear subspaces. We show that the concept of critical angles coincides with the concept of principal angles developed in the classic theory of angles between linear subspaces (e.g. [2]).

An important portion of this chapter is devoted to the analysis of critical angles in a pair of polyhedral cones. We provide a numerical method for computing all the critical angles for a given pair of polyhedral cones. We also give some estimations for the cardinality of the set of critical angles in this case.

Critical angles between two convex cones II. Special cases

This chapter focuses on the practical computation of the maximal angle and critical angles between specially structured cones. Revolution cones, topheavy cones and cone of matrices are in our list.

We start by considering a pair of revolution cones in any Euclidean space $\mathbb{X}$. In this case, we provide explicit formulas for computing each critical angle.

A topheavy cone in $\mathbb{R}^{n+1}$ is a closed convex cone of the form $\text{epi}f := \{(\xi, t) \in \mathbb{R}^{n+1} : f(\xi) \leq t\}$, where f is a norm on $\mathbb{R}^n$. The class of topheavy cones is quite large and includes in particular the ℓ^p- cones and the ellipsoidal cones. We show that under the assumption of "lower correlated norms" the maximal angle between two topheavy cones can be computed by considering the maximal angle of each cone (cf. [14]). This result is illustrated by computing the maximal angle between two ℓ^p- cones and between two ellipsoidal cones.

We have also obtained explicitly the maximal angle for the following situations: an ellipsoidal cone versus a nonnegative orthant and an ellipsoidal cone versus a ray.

Concerning the critical angles between two cones of symmetric matrices, we show that the set of critical angles between two "orthogonally invariant" cones is in correspondence with the set of critical angles between the respective permutation invariant cones.

At the end of this chapter we give a discussion about a difficult question arising in numerical linear algebra: how large can be the angle between a positive semidefinite symmetric matrix and a symmetric matrix that is nonnegative entrywise? We give some partial answers. For instance, we show that for dimension three the answer is $(3/4)\pi$ and for dimensions greater or equal than five the maximal angle must be greater than $(3/4)\pi$. By using numerical methods we have obtained lower bounds for this maximal angle for dimensions ranging from 4 to 30. Recently, Goldberg and Shaker-Monderer [4] have proved that this maximal angle tends to π when the dimension goes to ∞. It remains an open question to compute the exact value of this maximal angle for any dimension.

On the central path in symmetric cone linear programming

A symmetric cone linear program is an optimization problem of the form

$$(P) \quad \min_{x \in \mathbb{V}} \{\langle c, x\rangle : \mathcal{A}x = b,\ x \in \mathcal{K}\}$$

where $(\mathbb{V}, \circ, \langle \cdot, \cdot \rangle)$ is a Euclidean Jordan algebra, $\mathcal{K}$ is the cone of square elements on $\mathbb{V}$ (symmetric cone), $c \in \mathbb{V}$, $b \in \mathbb{R}^m$ and $\mathcal{A} : \mathbb{V} \to \mathbb{R}^m$ is a linear map. (P) is called the primal problem an its dual is denoted by (D).

The (primal-dual) central path is defined as the set $\{(x(\mu), y(\mu), s(\mu)) : \mu > 0\}$ derived from the optimality conditions of the penalized problem associated with (P), where the logarithm barrier function is used.

In this chapter we show that the central path converges to a point in the optimal sets of problems (P)-(D). By using the Peirce decomposition of $\mathbb{V}$ with respect to particular idempotents, we provide a characterization of the primal and dual optimal sets. We also conclude that the limit point is a maximally complementary solution and it lies in the relative interior of the primal and dual optimal sets.

A full characterization of the limit point is still an open problem. We expect to study this topic in a future work.

The results presented in this chapter are inspired by the results obtained for the central path in the particular context of semidefinite programming (cf. [5, 6, 9]).

Bibliography

[1] M. Baes. Convexity and differentiability properties of spectral functions and spectral mappings on Euclidean Jordan algebras. *Linear Algebra Appl.* 422 (2007), 664–700.

[2] A. Björck and G.H. Golub. Numerical methods for computing angles between linear subspaces. *Math. Comp.*, 27 (1973), 579–594.

[3] J. Faraut and A. Korányi. *Analysis on Symmetric Cones*. Clarendon Press, Oxford, 1994.

[4] F. Goldberg and N. Shaked-Monderer. On the maximal angle between copositive matrices. July 2013, submitted. Temporarily available at http://arxiv.org/pdf/1307.7519.pdf.

[5] M. Halická, E. De Klerk and C. Roos. On the convergence of the central path in semidefinite optimization. *SIAM J. Optim.* 12 (2002), 1090–1099.

[6] M. Halická, E. De Klerk and C. Roos. Limiting behavior of the central path in semidefinite optimization. *Optim. Methods Softw.* 20 (2005), 99–113.

[7] A. Iusem and A. Seeger. Angular analysis of two classes of non-polyhedral convex cones: the point of view of optimization theory. *Comput. Appl. Math.*, 26 (2007), 191–214.

[8] A. Iusem and A. Seeger. On pairs of vectors achieving the maximal angle of a convex cone. *Math. Program.*, 104 (2005), 501–523.

[9] E. de Klerk, C. Roos and T. Terlaky. Infeasible-start semidefinite programming algorithms via self-dual embeddings. *Fields Inst. Commun.,* 18 (1998), 215–236.

[10] M. Koecher. Jordan algebras and their applications. Lecture notes, Univ. of Minnesota, Minneapolis, 1962.

[11] A.S. Lewis. Convex analysis on the Hermitian matrices. *SIAM J. Optim.*, 6 (1996), 164–177.

[12] H. Ramírez, A. Seeger and D. Sossa. Commutation principle for variational problems on Euclidean Jordan algebras. *SIAM J. Optim.* 23 (2013), 687–694.

[13] H. Ramírez and D. Sossa. On the central path in symmetric cone linear programming. In preparation.

[14] A. Seeger. Epigraphical cones I. *J. Convex Analysis*, 18 (2011), 1171–1196.

[15] A. Seeger and D. Sossa. Complementarity problem with respect to Loewnerian cones. *J. Global Optim.*, 2014. Accepted.

[16] A. Seeger and D. Sossa. Critical angles between two convex cones I. General theory. March 2014, submitted.

[17] A. Seeger and D. Sossa. Critical angles between two convex cones II. Special cases. March 2014, submitted.

Chapter 1

Commutation principle for variational problems on Euclidean Jordan algebras[1]

Hector Ramírez[2] - Alberto Seeger[3] - David Sossa[4]

Abstract. This paper establishes a commutation result for variational problems involving spectral sets and spectral functions. The discussion takes places in the context of a general Euclidean Jordan algebra.

Mathematics Subject Classification. 15A18, 15A99, 17C99, 49J53.
Key words. Euclidean Jordan algebra, ordered Jordan frame, variational problem, operator commutation, spectral sets and spectral functions.

1.1 Introduction

Let the space $\mathbb{S}^n$ of real symmetric matrices of order n be equipped with the Frobenius or trace inner product $\langle X, Y\rangle = \mathrm{tr}(XY)$. The notation $\lambda(X)$ refers to the (column) vector of eigenvalues

$$\lambda_1(X) \leq \ldots \leq \lambda_n(X)$$

[1] A shortened version of this chapter was published on *SIAM Journal on Optimization*, see [13].

[2]Departamento de Ingeniería Matemática, Centro de Modelamiento Matemático (CNRS UMI 2807), FCFM, Universidad de Chile, Blanco Encalada 2120, Santiago, Chile (E-mail: hramirez@dim.uchile.cl). This author is supported by FONDECYT project No. 1110888 and BASAL Project (Centro de Modelamiento Matemático, Universidad de Chile).

[3]University of Avignon, Department of Mathematics, 33 rue Louis Pasteur, 84000 Avignon, France (E-mail: alberto.seeger@univ-avignon.fr).

[4]Departamento de Ingeniería Matemática, Centro de Modelamiento Matemático (CNRS UMI 2807), FCFM, Universidad de Chile, Blanco Encalada 2120, Santiago, Chile (E-mail: dsossa@dim.uchile.cl). This author is supported by CONICYT (Chile).

of $X \in \mathbb{S}^n$ arranged in nondecreasing order. Recall that a spectral set in $\mathbb{S}^n$ is a set of the form

$$\Omega = \lambda^{-1}(Q) := \{X \in \mathbb{S}^n : \lambda(X) \in Q\},$$

where Q is a permutation invariant set in $\mathbb{R}^n$. A spectral function on $\mathbb{S}^n$ is a real-valued function $\Phi : \mathbb{S}^n \to \mathbb{R}$ admitting the representation

$$\Phi(X) = g(\lambda(X)),$$

where $g : \mathbb{R}^n \to \mathbb{R}$ is a real-valued permutation invariant function. Spectrality is a property that some authors refer to as orthogonal invariance. General information on the theory of spectral sets and spectral functions can be found in [3, 11, 14] and the references therein.

Iusem and Seeger [9, Lemma 4] established recently the following commutativity result for variational problems involving spectral data. By a local extremum of a function one understands a local minimum or a local maximum.

Lemma 1.1. *Let $A, B \in \mathbb{S}^n$. Suppose that $\Omega \subseteq \mathbb{S}^n$ is a spectral set and that $\Phi : \mathbb{S}^n \to \mathbb{R}$ is a spectral function. Under these assumptions, if B is a local extremum of*

$$X \in \Omega \mapsto F(X) = \langle A, X\rangle + \Phi(X),$$

then A and B commute, i.e., $AB = BA$.

It is worthwhile to keep in mind that if two symmetric matrices commute, then it is possible to diagonalize them by means of a common orthogonal matrix. The possibility of simultaneous diagonalization opens the way to significant simplifications in the proof of various linear algebra results. The commutation principle stated in Lemma 1.1 has applications in various fields, see for instance

- Fenchel conjugate and subdifferential of a convex spectral function (cf.[11]),
- distance to a spectral set (cf. [3, Proposition 2.3]),
- inradius and incenter of a spectral convex cone (cf. [7, Theorem 3.3]),
- distance between a pair of spectral convex cones (cf. [10, Proposition 6.6]),
- antipodal pairs in spectral convex cones (cf. [9, Theorem 4]).

It turns out that Lemma 1.1 is a particular instance of a more general and deep commutation principle for variational problems on Euclidean Jordan algebras. The main result of this paper reads as follows:

Theorem 1.2. *Let $(\mathbb{V}, \circ, \langle\cdot,\cdot\rangle)$ be a Euclidean Jordan algebra and let $a, b \in \mathbb{V}$. Suppose that $\Omega \subseteq \mathbb{V}$ is a spectral set and that $\Phi : \mathbb{V} \to \mathbb{R}$ is a spectral function. Under these assumptions, if b is a local extremum of*

$$x \in \Omega \mapsto F(x) = \langle a, x\rangle + \Phi(x), \tag{1.1}$$

then a and b operator commute, i.e.,

$$a \circ (b \circ z) = b \circ (a \circ z) \quad \text{for all } z \in \mathbb{V}. \tag{1.2}$$

For simplicity in the exposition we consider Φ as a spectral function on the whole space $\mathbb{V}$, but one could restrict Φ to the spectral subset Ω. Section 1.2 reviews some basic material on Euclidean Jordan algebras and prepares the ground for proving Theorem 1.2. The proof itself is given in Section 1.3. Some applications are mentioned in Section 1.4.

1.2 Preliminary material on Euclidean Jordan algebras

Throughout this work one assumes that $(\mathbb{V}, \circ, \langle\cdot,\cdot\rangle)$ is a Euclidean Jordan algebra (EJA) with unit element $e \in \mathbb{V}$. This means that $\mathbb{V}$ is a finite dimensional real vector space equipped with an inner product $\langle\cdot,\cdot\rangle$ and a bilinear function $\circ : \mathbb{V} \times \mathbb{V} \to \mathbb{V}$ satisfying the axioms:

$$\begin{cases} x \circ y = y \circ x & \text{for all } x, y \in \mathbb{V}, \\ x \circ (x^2 \circ y) = x^2 \circ (x \circ y) & \text{for all } x, y \in \mathbb{V}, \\ \langle x \circ y, z\rangle = \langle y, x \circ z\rangle & \text{for all } x, y, z \in \mathbb{V}, \\ e \circ x = x & \text{for all } x \in \mathbb{V}. \end{cases}$$

The unit element e is clearly unique. Here and in the sequel one uses the notation $x^2 = x \circ x$. Higher order powers are defined recursively by $x^{k+1} = x \circ x^k$. The rank of $\mathbb{V}$ is declared to be

$$r = \max\{\deg(x) : x \in \mathbb{V}\},$$

where $\deg(x)$ is the smallest positive integer k such that $\{e, x, x^2, \ldots, x^k\}$ is linearly dependent.

The Lyapunov operator associated to a given $x \in \mathbb{V}$ is the linear map $L_x : \mathbb{V} \to \mathbb{V}$ given by $L_x y = x \circ y$. The operator commutation property (1.2) amounts to saying that the bracket

$$[L_a, L_b] := L_a L_b - L_b L_a$$

is equal to the zero map on $\mathbb{V}$.

An element $c \in \mathbb{V}$ is an idempotent if $c^2 = c$. An idempotent c is primitive if it is nonzero and cannot be written as a sum of two nonzero idempotents. A Jordan frame is a collection $\{c_1, \ldots, c_r\}$ of primitive idempotents satisfying

$$\sum_{i=1}^{r} c_i = e \quad \text{and} \quad c_i \circ c_j = 0 \ \text{ when } i \neq j.$$

We recall below a spectral decomposition theorem taken from [4, Theorem III.1.2].

Theorem 1.3. *Let $(\mathbb{V}, \circ, \langle\cdot,\cdot\rangle)$ be an EJA with rank r. Then, for every $x \in \mathbb{V}$, there exists a Jordan frame $\{c_1, \ldots, c_r\}$ and real numbers $\lambda_1, \ldots, \lambda_r$ such that*

$$x = \lambda_1 c_1 + \ldots + \lambda_r c_r.$$

The λ_i's are uniquely determined by x.

We write $\lambda_i(x)$ to underline the dependence with respect to x. Renumbering the c_i's if necessary, one may suppose that the $\lambda_i(x)$'s are arranged in nondecreasing order, i.e.,

$$\lambda_1(x) \leq \ldots \leq \lambda_r(x).$$

By analogy with the case of symmetric matrices, one refers to the (column) vector $\lambda(x) \in \mathbb{R}^r$ as the vector of "eigenvalues" of $x \in \mathbb{V}$. A spectral set in $\mathbb{V}$ is then a set of the form

$$\Omega = \lambda^{-1}(Q) := \{x \in \mathbb{V} : \lambda(x) \in Q\} \tag{1.3}$$

with $Q \subseteq \mathbb{R}^r$ permutation invariant. A spectral function on $\mathbb{V}$ is a real-valued function $\Phi : \mathbb{V} \to \mathbb{R}$ admitting the representation

$$\Phi(x) = g(\lambda(x))$$

with $g : \mathbb{R}^r \to \mathbb{R}$ permutation invariant. The formulation of Theorem 1.2 is now perfectly clear.

Remark 1.4. Note that Lemma 1.1 can be derived from Theorem 1.2 by working with the particular EJA

$$\begin{cases} \mathbb{V} = \mathbb{S}^n, \\ \langle X, Y \rangle = \mathrm{tr}(XY), \\ X \circ Y = \frac{1}{2}(XY + YX), \\ e = I_n \text{ (identity matrix of order } n). \end{cases} \tag{1.4}$$

One can easily check that the operator commutation property

$$A \circ (B \circ Z) = B \circ (A \circ Z) \quad \text{for all } Z \in \mathbb{S}^n$$

is equivalent to the usual commutation condition $AB = BA$.

The following result, borrowed from [1, Theorem 27], shows the importance of the concept of operator commutation. We mention in passing that this concept admits also other equivalent characterizations, see for instance in [12, Theorem 1].

Theorem 1.5. *Let $(\mathbb{V}, \circ, \langle \cdot, \cdot \rangle)$ be an EJA with rank r. Then $a \in \mathbb{V}$ and $b \in \mathbb{V}$ operator commute if and only if a and b admit a common Jordan frame, i.e., there exist a Jordan frame $\{c_1, \ldots, c_r\}$ and real numbers $\lambda_1, \ldots, \lambda_r$ and $\mu_1, \ldots, \mu_r$ such that*

$$\begin{aligned} a &= \lambda_1 c_1 + \ldots + \lambda_r c_r \\ b &= \mu_1 c_1 + \ldots + \mu_r c_r. \end{aligned}$$

1.2.1 The tangent space to the set of ordered Jordan frames

The proof of Theorem 1.2 relies on the analysis of an optimization problem of the form

$$\begin{cases} \text{minimize } f(\mathbf{c}) \\ \mathbf{c} \in \mathcal{O}_{\mathbb{V}}, \end{cases} \tag{1.5}$$

where $f : \mathbb{V}^r \to \mathbb{R}$ is a continuously differentiable function and

$$\mathcal{O}_{\mathbb{V}} := \{\mathbf{c} = (c_1, \ldots, c_r) \in \mathbb{V}^r : \{c_1, \ldots, c_r\} \text{ is a Jordan frame}\}. \tag{1.6}$$

Each element of (1.6) is called an ordered Jordan frame. A local solution $\bar{\mathbf{c}}$ to the problem (1.5) satisfies the first-order optimality condition

$$f'(\bar{\mathbf{c}})\mathbf{h} \geq 0 \quad \text{for all } \mathbf{h} \in T_{\bar{\mathbf{c}}}[\mathcal{O}_{\mathbb{V}}], \tag{1.7}$$

where $f'(\bar{\mathbf{c}}) : \mathbb{V}^r \to \mathbb{R}$ is the differential map of f at $\bar{\mathbf{c}}$ and $T_{\bar{\mathbf{c}}}[\mathcal{O}_{\mathbb{V}}]$ is the Bouligand tangent set to $\mathcal{O}_{\mathbb{V}}$ at $\bar{\mathbf{c}}$ (cf. [2, Definition 4.1.1]).

The next lemma shows that $T_{\bar{\mathbf{c}}}[\mathcal{O}_{\mathbb{V}}]$ is a linear subspace and provides an explicit formula for computing this set. It also characterizes the orthogonal complement

$$(T_{\bar{\mathbf{c}}}[\mathcal{O}_{\mathbb{V}}])^{\perp} := \left\{ \mathbf{q} \in \mathbb{V}^r : \sum_{i=1}^{r} \langle q_i, h_i \rangle = 0 \text{ for all } \mathbf{h} \in T_{\bar{\mathbf{c}}}[\mathcal{O}_{\mathbb{V}}] \right\}.$$

For notational convenience we introduce the index sets

$$\begin{aligned} N_r &:= \{1, \ldots, r\} \\ M_r &:= \{(i,j) \in N_r \times N_r : i < j\}. \end{aligned}$$

Lemma 1.6. *Let $(\mathbb{V}, \circ, \langle \cdot, \cdot \rangle)$ be an EJA with rank r and let $\bar{\mathbf{c}} \in \mathcal{O}_{\mathbb{V}}$. Then the following hold*

(a) $\mathbf{h} = (h_1, \ldots, h_r) \in \mathbb{V}^r$ *belongs to* $T_{\bar{\mathbf{c}}}[\mathcal{O}_{\mathbb{V}}]$ *if and only if*

$$2\bar{c}_i \circ h_i - h_i = 0 \quad \text{for all } i \in N_r \tag{1.8}$$
$$\bar{c}_i \circ h_j + \bar{c}_j \circ h_i = 0 \quad \text{for all } (i,j) \in M_r. \tag{1.9}$$

In particular, $T_{\bar{\mathbf{c}}}[\mathcal{O}_{\mathbb{V}}]$ *is a linear subspace.*

(b) $\mathbf{q} = (q_1, \ldots, q_r) \in \mathbb{V}^r$ *belongs to* $(T_{\bar{\mathbf{c}}}[\mathcal{O}_{\mathbb{V}}])^{\perp}$ *if and only if there are vectors*

$$\{\alpha_i : i \in N_r\} \subseteq \mathbb{V} \quad \text{and} \quad \{\beta_{i,j} : (i,j) \in M_r\} \subseteq \mathbb{V} \tag{1.10}$$

such that

$$q_i = 2\alpha_i \circ \bar{c}_i - \alpha_i + \sum_{j=1}^{i-1} \beta_{j,i} \circ \bar{c}_j + \sum_{j=i+1}^{r} \beta_{i,j} \circ \bar{c}_j \quad \text{for all } i \in N_r. \tag{1.11}$$

Proof. Part (a). Let $\mathcal{R}_{\mathbb{V}}$ be the set of all $\mathbf{c} \in \mathbb{V}^r$ satisfying the nonlinear system

$$\begin{cases} c_i^2 - c_i = 0 & \text{for all } i \in N_r \\ c_i \circ c_j = 0 & \text{for all } (i,j) \in M_r. \end{cases} \tag{1.12}$$

Since $\mathcal{O}_{\mathbb{V}} \subseteq \mathcal{R}_{\mathbb{V}}$, it is clear that

$$T_{\bar{\mathbf{c}}}[\mathcal{O}_{\mathbb{V}}] \subseteq T_{\bar{\mathbf{c}}}[\mathcal{R}_{\mathbb{V}}].$$

Let $\mathcal{B}_{\bar{\mathbf{c}}}$ be the set of all $\mathbf{h} \in \mathbb{V}^r$ satisfying the linear system (1.8)-(1.9). Observe that (1.8)-(1.9) is obtained by linearizing (1.12) around the reference point $\bar{\mathbf{c}}$. This observation shows that

$$T_{\bar{\mathbf{c}}}[\mathcal{R}_{\mathbb{V}}] \subseteq \mathcal{B}_{\bar{\mathbf{c}}}.$$

For completing the proof of (a) it suffices to check that

$$\mathcal{B}_{\bar{\mathbf{c}}} \subseteq T_{\bar{\mathbf{c}}}[\mathcal{O}_{\mathbb{V}}].$$

Let $\mathbf{h} \in \mathcal{B}_{\bar{\mathbf{c}}}$. To start with, we check that

$$\sum_{i=1}^{r} h_i = 0. \tag{1.13}$$

One has

$$\begin{aligned} \sigma &:= \sum_{i=1}^{r}\sum_{j=1}^{r} \{\bar{c}_i \circ h_j + \bar{c}_j \circ h_i\} \\ &= \left(\sum_{i=1}^{r} \bar{c}_i\right) \circ \left(\sum_{j=1}^{r} h_j\right) + \left(\sum_{j=1}^{r} \bar{c}_j\right) \circ \left(\sum_{i=1}^{r} h_i\right) \\ &= 2\sum_{i=1}^{r} h_i. \end{aligned}$$

On the other hand,

$$\sigma = \sum_{i=1}^{r} 2\,\bar{c}_i \circ h_i + 2 \sum_{(i,j)\in M_r} \{\bar{c}_i \circ h_j + \bar{c}_j \circ h_i\} = \sum_{i=1}^{r} h_i.$$

This confirms the equality (1.13). Next, we construct a continuously differentiable function

$$t \in \mathbb{R} \mapsto \gamma(t) = (\gamma_1(t), \ldots, \gamma_r(t))$$

such that

$$\gamma(t) \in \mathcal{O}_{\mathbb{V}} \quad \text{for all } t \in \mathbb{R} \tag{1.14}$$
$$\gamma(0) = \bar{\mathbf{c}} \tag{1.15}$$
$$\gamma'(0) = \mathbf{h}. \tag{1.16}$$

The existence of such function γ implies that $\mathbf{h} \in T_{\bar{\mathbf{c}}}[\mathcal{O}_{\mathbb{V}}]$. We suggest to take as the i-th component of γ a function of the form

$$\gamma_i(t) = \underbrace{\exp(tD)}_{A_t}\, \bar{c}_i,$$

where $D : \mathbb{V} \to \mathbb{V}$ is some linear map and $\{A_t\}_{t\in\mathbb{R}}$ is the associated semigroup, i.e.,

$$A_t := \sum_{p=0}^{\infty} \frac{t^p}{p!}\, D^p.$$

The condition (1.15) clearly holds and $\gamma'(0) = (D\bar{c}_1, \ldots, D\bar{c}_r)$. Hence, D is to be constructed so that

$$D\bar{c}_i = h_i \quad \text{for all} \quad i \in N_r. \tag{1.17}$$

Inspired in the technique used in [8, Section 3], we take

$$D := 2\sum_{k=1}^{r}[L_{h_k}, L_{\bar{c}_k}]$$

is a suitable choice. A matter of computation shows that

$$\begin{aligned} D\bar{c}_i &= 2\sum_{k=1}^{r} h_k \circ (\bar{c}_k \circ \bar{c}_i) - 2\sum_{k=1}^{r} \bar{c}_k \circ (h_k \circ \bar{c}_i) \\ &= 2h_i \circ (\bar{c}_i \circ \bar{c}_i) - 2\bar{c}_i \circ (h_i \circ \bar{c}_i) - 2\sum_{k\in N_r\setminus\{i\}} \bar{c}_k \circ (h_k \circ \bar{c}_i). \end{aligned}$$

Since $\bar{c}_i$ is idempotent and (1.8) holds, one gets

$$\begin{aligned} D\bar{c}_i &= \frac{1}{2}h_i - 2\sum_{k\in N_r\setminus\{i\}} \bar{c}_k \circ (h_k \circ \bar{c}_i) \\ &= \frac{1}{2}h_i - 2\sum_{k\in N_r\setminus\{i\}} \bar{c}_i \circ (h_k \circ \bar{c}_k), \end{aligned}$$

the last equality being due to the fact that $\bar{c}_k$ and $\bar{c}_i$ operator commute (cf. [5, Proposition 6]). By using (1.8) and simplifying, one obtains

$$\begin{aligned} D\bar{c}_i &= \frac{1}{2}h_i - \sum_{k\in N_r\setminus\{i\}} \bar{c}_i \circ h_k \\ &= \frac{1}{2}h_i - \bar{c}_i \circ \left(\sum_{k\in N_r\setminus\{i\}} h_k\right). \end{aligned}$$

By using (1.13) and (1.8), one gets finally

$$D\bar{c}_i = \frac{1}{2}h_i - \bar{c}_i \circ (-h_i) = h_i.$$

This proves (1.17) and the condition (1.16). We now take care of (1.14), that is to say, we pick an arbitrary $t \in \mathbb{R}$ and check that $\{\gamma_1(t), \ldots, \gamma_r(t)\}$ is a Jordan frame. Thanks to [4, Proposition II.4.1] one knows that D is a derivation on $\mathbb{V}$, i.e.,

$$D(x \circ y) = (Dx) \circ y + x \circ Dy \quad \text{for all } x, y \in \mathbb{V}.$$

Such a property implies in turn that $A_t : \mathbb{V} \to \mathbb{V}$ satisfies the identity

$$(A_t x) \circ (A_t y) = A_t(x \circ y) \quad \text{for all } x, y \in \mathbb{V}.$$

In particular, one has

$$\gamma_i(t) \circ \gamma_j(t) = (A_t\bar{c}_i) \circ (A_t\bar{c}_j) = A_t(\bar{c}_i \circ \bar{c}_j) = 0$$

whenever i and j are different. Similarly,

$$\gamma_i(t) \circ \gamma_i(t) = (A_t\bar{c}_i) \circ (A_t\bar{c}_i) = A_t(\bar{c}_i \circ \bar{c}_i) = A_t\bar{c}_i = \gamma_i(t),$$

that is to say, $\gamma_i(t)$ is idempotent. We claim that $\gamma_i(t)$ is also primitive. Indeed, if $\gamma_i(t)$ is not primitive, then one can write

$$\gamma_i(t) = A_t\bar{c}_i = \bar{a}_i + \bar{b}_i,$$

where a_i and b_i are nonzero idempotents. Hence,

$$\bar{c}_i = A_t^{-1}\bar{a}_i + A_t^{-1}\bar{b}_i.$$

Since $-D$ is a derivation and $\{A_t^{-1}\}_{t\in\mathbb{R}}$ is its associated semigroup, it follows that $A_t^{-1}\bar{a}_i$ and $A_t^{-1}\bar{b}_i$ are nonzero idempotents, contradicting the fact that $\bar{c}_i$ is primitive. Finally, we observe that

$$\sum_{i=1}^{r} \gamma_i(t) = \sum_{i=1}^{r} A_t\bar{c}_i = A_t\left(\sum_{i=1}^{r} \bar{c}_i\right) = A_t e.$$

Keeping in mind that D is a derivation, one sees that

$$De = D(e \circ e) = (De) \circ e + e \circ (De) = 2\, De.$$

Thus $De = 0$ and $A_t e = e$. This completes the proof of (a).

Part (b). The linear system (1.8)-(1.9) is formed by

$$s_r := r + \frac{r(r-1)}{2}$$

equations. So, the linear subspace $\mathcal{B}_{\bar{\mathbf{c}}}$ corresponds to the kernel of a linear map $\mathcal{M} : \mathbb{V}^r \to \mathbb{V}^{s_r}$ whose definition is clear. Hence, the orthogonal complement of $\mathcal{B}_{\bar{\mathbf{c}}}$ is equal to the range of the adjoint map $\mathcal{M}^* : \mathbb{V}^{s_r} \to \mathbb{V}^r$. The explicit computation of this adjoint leads to the announced formula (1.11). The details are omitted. □

1.3 Proof of the general commutation principle

This section takes care of the proof of Theorem 1.2. In fact, without extra effort one can demonstrate a generalized version of Theorem 1.2 in which the linear functional $\langle a, \cdot\rangle$ is changed by a nonlinear differentiable function $\mathcal{E} : \mathbb{V} \to \mathbb{R}$.

Theorem 1.7. *Let $(\mathbb{V}, \circ, \langle\cdot,\cdot\rangle)$ be a Euclidean Jordan algebra and let $b \in \mathbb{V}$ be a point at which $\mathcal{E} : \mathbb{V} \to \mathbb{R}$ is continuously differentiable. Suppose that $\Omega \subseteq \mathbb{V}$ is a spectral set and that $\Phi : \mathbb{V} \to \mathbb{R}$ is a spectral function. Under these assumptions, if b is a local extremum of*

$$x \in \Omega \mapsto F(x) = \mathcal{E}(x) + \Phi(x), \tag{1.18}$$

then b and $\nabla\mathcal{E}(b)$ operator commute.

Proof. Suppose that b is a local minimum of (1.18), the case of a local maximum can be treated in a similar way. One has $b \in \Omega$ and

$$\mathcal{E}(b) + \Phi(b) \leq \mathcal{E}(x) + \Phi(x) \quad \text{for all } x \in \Omega \cap \mathcal{N}_b,$$

where $\mathcal{N}_b$ is some neighborhood of b. Let $\bar{\mathbf{c}} \in \mathcal{O}_{\mathbb{V}}$ be an ordered Jordan frame such that

$$b = \sum_{i=1}^{r} \bar{\lambda}_i \bar{c}_i$$

with $\bar{\lambda}_i = \lambda_i(b)$. Consider the linear function $\Gamma : \mathbb{V}^r \to \mathbb{V}$ defined by

$$\Gamma(\mathbf{c}) = \sum_{i=1}^{r} \bar{\lambda}_i c_i.$$

Let $x := \Gamma(\mathbf{c})$ with $\mathbf{c} \in \mathbb{V}^r$. If $\mathbf{c}$ is taken in a small neighborhood $\mathcal{N}_{\bar{\mathbf{c}}}$ of $\bar{\mathbf{c}}$, then $x \in \mathcal{N}_b$ by the continuity of Γ. On the other hand, if $\mathbf{c}$ belongs to $\mathcal{O}_{\mathbb{V}}$, then $\lambda(x) = \lambda(b)$ and, a posteriori, $x \in \Omega$. Hence,

$$\mathcal{E}\left(\sum_{i=1}^{r} \bar{\lambda}_i \bar{c}_i\right) + \Phi\left(\sum_{i=1}^{r} \bar{\lambda}_i \bar{c}_i\right) \leq \mathcal{E}\left(\sum_{i=1}^{r} \bar{\lambda}_i c_i\right) + \Phi\left(\sum_{i=1}^{r} \bar{\lambda}_i c_i\right) \quad \text{for all } \mathbf{c} \in \mathcal{O}_{\mathbb{V}} \cap \mathcal{N}_{\bar{\mathbf{c}}}.$$

The spectrality of Φ leads to the simpler inequality

$$\mathcal{E}\left(\sum_{i=1}^{r} \bar{\lambda}_i \bar{c}_i\right) \leq \mathcal{E}\left(\sum_{i=1}^{r} \bar{\lambda}_i c_i\right) \quad \text{for all } \mathbf{c} \in \mathcal{O}_{\mathbb{V}} \cap \mathcal{N}_{\bar{\mathbf{c}}}.$$

We have shown in this way that $\bar{\mathbf{c}}$ is a local minimum on $\mathcal{O}_{\mathbb{V}}$ of the function

$$\mathbf{c} \in \mathbb{V}^r \mapsto f(\mathbf{c}) = \mathcal{E}\left(\sum_{i=1}^{r} \bar{\lambda}_i c_i\right).$$

Note that f is differentiable at $\bar{\mathbf{c}}$ because $\mathcal{E}$ is differentiable at b. The optimality condition (1.7) takes the particular form

$$\underbrace{(\bar{\lambda}_1 a, \ldots, \bar{\lambda}_r a)}_{\nabla f(\bar{\mathbf{c}})} \in (T_{\bar{\mathbf{c}}}[\mathcal{O}_{\mathbb{V}}])^{\perp},$$

where $a := \nabla \mathcal{E}(b)$. In view of the characterization (1.11) of the subspace $(T_{\bar{\mathbf{c}}}[\mathcal{O}_{\mathbb{V}}])^{\perp}$, one gets

$$\bar{\lambda}_i a = 2\alpha_i \circ \bar{c}_i - \alpha_i + \sum_{j=1}^{i-1} \beta_{j,i} \circ \bar{c}_j + \sum_{j=i+1}^{r} \beta_{i,j} \circ \bar{c}_j \quad \text{for all } i \in N_r \tag{1.19}$$

for suitable vectors α_i and $\beta_{i,j}$ as in (1.10). With this information at hand, we are now ready to show that a and b operator commute. One has

$$\begin{aligned}
[L_a, L_b] &= L_a \left(\sum_{i=1}^{r} \bar{\lambda}_i L_{\bar{c}_i}\right) - \left(\sum_{i=1}^{r} \bar{\lambda}_i L_{\bar{c}_i}\right) L_a \\
&= \left(\sum_{i=1}^{r} L_{\bar{\lambda}_i a} L_{\bar{c}_i}\right) - \left(\sum_{i=1}^{r} L_{\bar{c}_i} L_{\bar{\lambda}_i a}\right) \\
&= \sum_{i=1}^{r} [L_{\bar{\lambda}_i a}, L_{\bar{c}_i}].
\end{aligned}$$

But (1.19) implies that

$$L_{\bar{\lambda}_i a} = 2L_{\alpha_i \circ \bar{c}_i} - L_{\alpha_i} + L_{\nu_i}$$

with

$$\nu_i := \sum_{j=1}^{i-1} \beta_{j,i} \circ \bar{c}_j + \sum_{j=i+1}^{r} \beta_{i,j} \circ \bar{c}_j.$$

Hence

$$[L_a, L_b] = \underbrace{\sum_{i=1}^{r} \left\{ 2[L_{\alpha_i \circ \bar{c}_i}, L_{\bar{c}_i}] - [L_{\alpha_i}, L_{\bar{c}_i}] \right\}}_{\Delta_1} + \underbrace{\sum_{i=1}^{r} [L_{\nu_i}, L_{\bar{c}_i}]}_{\Delta_2}.$$

We claim that Δ_1 is the zero map on $\mathbb{V}$. Indeed,

$$2[L_{\alpha_i \circ \bar{c}_i}, L_{\bar{c}_i}] = [L_{\alpha_i}, L_{\bar{c}_i}] \quad \text{for all} \quad i \in N_r,$$

as one can see from the general identity (cf. [4, Proposition II.1.1])

$$2[L_{u \circ z}, L_z] = [L_u, L_{z^2}] \quad \text{for all} \quad u, z \in \mathbb{V}$$

and the fact that $\bar{c}_i$ is idempotent. Also Δ_2 is the zero map on $\mathbb{V}$. To see this, we write

$$\begin{aligned}
\Delta_2 &= \sum_{i=1}^{r} \left[\sum_{j=1}^{i-1} L_{\beta_{j,i} \circ \bar{c}_j} + \sum_{j=i+1}^{r} L_{\beta_{i,j} \circ \bar{c}_j}, L_{\bar{c}_i} \right] \\
&= \sum_{i=1}^{r} \left\{ \sum_{j=1}^{i-1} \left[L_{\beta_{j,i} \circ \bar{c}_j}, L_{\bar{c}_i} \right] + \sum_{j=i+1}^{r} \left[L_{\beta_{i,j} \circ \bar{c}_j}, L_{\bar{c}_i} \right] \right\} \\
&= \sum_{(i,j) \in M_r} \left\{ \left[L_{\beta_{i,j} \circ \bar{c}_i}, L_{\bar{c}_j} \right] + \left[L_{\beta_{i,j} \circ \bar{c}_j}, L_{\bar{c}_i} \right] \right\}
\end{aligned}$$

and observe that

$$\left[L_{\beta_{i,j} \circ \bar{c}_i}, L_{\bar{c}_j} \right] + \left[L_{\beta_{i,j} \circ \bar{c}_j}, L_{\bar{c}_i} \right] = 0.$$

The above equality follows from the general identity (cf. [4, Proposition II.1.1])

$$[L_{u \circ y}, L_z] + [L_{u \circ z}, L_y] = [L_u, L_{z \circ y}] \quad \text{for all} \quad u, y, z \in \mathbb{V}$$

and the fact that $\bar{c}_i \circ \bar{c}_j = 0$ for all $(i, j) \in M_r$. This shows that a and b operator commute, finishing the proof. □

Remark 1.8. A special case of Theorem 1.7 is obtained when Φ is the zero function. It reads as follows: If b is local extremum of

$$x \in \Omega \mapsto \mathcal{E}(x),$$

then b and $\nabla \mathcal{E}(b)$ operator commute.

1.4 Applications

Some simple but illuminating examples suffice to illustrate how the commutation principle works in practice.

1.4.1 Variational inequalities

The next proposition concerns a variational inequality of the form

$$\begin{cases} \langle G(x), u - x\rangle + \Phi(u) - \Phi(x) \geq 0 \ \text{ for all } \ u \in \Omega \\ x \in \Omega, \end{cases} \tag{1.20}$$

where $G : \mathbb{V} \to \mathbb{V}$ is an arbitrary function.

Proposition 1.9. *Let $(\mathbb{V}, \circ, \langle\cdot,\cdot\rangle)$ be an EJA. Suppose that $\Omega \subseteq \mathbb{V}$ is a spectral set and that $\Phi : \mathbb{V} \to \mathbb{R}$ is a spectral function. Under these assumptions, if b is a solution to the variational inequality (1.20), then b and $G(b)$ operator commute.*

Proof. If b solves (1.20), then $b \in \Omega$ and

$$\langle G(b), u\rangle + \Phi(u) \ \geq \ \langle G(b), b\rangle + \Phi(b) \ \text{ for all } \ u \in \Omega.$$

Hence, b is a global minimum of

$$u \in \Omega \mapsto F(u) = \langle G(b), u\rangle + \Phi(u).$$

Theorem 1.2 leads to the announced conclusion. □

1.4.2 Distance to a spectral set

Recall that the trace operator on a Euclidean Jordan algebra $(\mathbb{V}, \circ, \langle\cdot,\cdot\rangle)$ with rank r is the real-valued function

$$u \in \mathbb{V} \mapsto \operatorname{Tr}(u) = \sum_{i=1}^{r} \lambda_i(u).$$

An EJA is said to be *scalarizable* if there exists a positive constant θ such that

$$\langle x, y\rangle = \theta \operatorname{Tr}(x \circ y) \quad \text{ for all } \ x, y \in \mathbb{V}. \tag{1.21}$$

Such constant is unique and called the scaling factor of the EJA. Of course, it is given by

$$\theta \ = \ \frac{\langle e, e\rangle}{\operatorname{Tr}(e)}.$$

As a consequence of the scalarization property (1.21) one gets

$$\|x\| = \sqrt{\theta}\, \|\lambda(x)\|_2 \quad \text{ for all } \ x \in \mathbb{V}, \tag{1.22}$$

where $\|\cdot\|_2$ is the usual Euclidean norm on $\mathbb{R}^r$. It is well known that:

- The EJA given by (1.4) is of rank $r = n$. It is scalarizable and has $\theta = 1$ as scaling factor.

- The EJA given by
$$\begin{cases} \mathbb{V} = \mathbb{R}^{n-1} \times \mathbb{R}, \\ \langle (\xi, t), (\eta, s) \rangle = \xi^T \eta + ts, \\ (\xi, t) \circ (\eta, s) = \left(s\xi + t\eta, \xi^T \eta + ts \right), \\ e = ((0, \ldots, 0), 1)^T \end{cases}$$
is of rank $r = 2$. It is scalarizable and has $\theta = 1/2$ as scaling factor.

Remark 1.10. If an EJA is *simple* in the sense that it does not contain any nontrivial ideal, then it is scalarizable, see [4, Proposition III.4.1].

The next proposition provides a formula for computing the distance

$$\operatorname{dist}[a, \Omega] := \inf_{x \in \Omega} \|a - x\| \tag{1.23}$$

from a point $a \in \mathbb{V}$ to a spectral set $\Omega \subseteq \mathbb{V}$. Our result is a generalization of [3, Proposition 2.3].

Proposition 1.11. *Let $(\mathbb{V}, \circ, \langle \cdot, \cdot \rangle)$ be an EJA with rank r and scaling factor θ. Let $\Omega \subseteq \mathbb{V}$ be a spectral set as in (1.3). Then for all $a \in \mathbb{V}$ one has*

$$\frac{1}{\sqrt{\theta}} \operatorname{dist}[a, \Omega] \;=\; \inf_{\mu \in Q} \|\lambda(a) - \mu\|_2. \tag{1.24}$$

In particular, $\operatorname{dist}[\,\cdot\,, \Omega]$ *is a spectral function.*

Proof. The topological closure of Q is permutation invariant and one has

$$\operatorname{cl}(\Omega) = \lambda^{-1}(\operatorname{cl}(Q)).$$

Hence, $\operatorname{cl}(\Omega)$ is a spectral set. Since a distance function is blind with respect to topological closure, there is no loss of generality in assuming that Q is already closed (in which case, also Ω is closed). Let b be a solution to (1.23). Then b is a global maximum of

$$x \in \Omega \mapsto F(x) = \langle a, x \rangle - \frac{1}{2}\|x\|^2.$$

But (1.22) implies that $-(1/2)\|\cdot\|^2$ is a spectral function on $\mathbb{V}$. Hence, a and b operator commute by Theorem 1.2. Thanks to Theorem 1.5, there exists $\mathbf{c} \in \mathcal{O}_{\mathbb{V}}$ and $\bar{\mu} \in \mathbb{R}^r$ such that

$$a = \sum_{i=1}^{r} \lambda_i(a) c_i \quad \text{and} \quad b = \sum_{i=1}^{r} \bar{\mu}_i c_i.$$

Clearly,

$$\operatorname{dist}[a, \Omega] = \|a - b\| = \sqrt{\theta}\, \|\lambda(a) - \bar{\mu}\|_2.$$

We claim that $\bar{\mu}$ solves the minimization problem on the right-hand side of (1.24). Up to a permutation, the vector $\bar{\mu}$ is equal to $\lambda(b)$. Hence, $\bar{\mu} \in Q$. Suppose that there exists $\tilde{\mu} \in Q$ such that

$$\|\lambda(a) - \tilde{\mu}\|_2 < \|\lambda(a) - \bar{\mu}\|_2.$$

In such a case $\tilde{x} = \sum_{i=1}^{r} \tilde{\mu}_i c_i$ belongs to Ω and

$$\|a - \tilde{x}\| \;=\; \sqrt{\theta}\, \|\lambda(a) - \tilde{\mu}\|_2 \;<\; \sqrt{\theta}\, \|\lambda(a) - \bar{\mu}\|_2 \;=\; \operatorname{dist}[a, \Omega],$$

a clear contradiction. □

1.4.3 Inradius of a spectral cone

Proposition 1.11 has in turn several applications. For instance, it can be used to derive a formula for computing the inradius

$$\rho(K) := \sup_{\substack{x \in K \\ \|x\|=1}} \operatorname{dist}[x, \operatorname{bd}(K)] \tag{1.25}$$

of a spectral proper cone K in $\mathbb{V}$. By a proper cone one understands a closed convex cone that is pointed and has nonempty interior. The notation $\operatorname{bd}(K)$ refers to the boundary of K. The maximization problem (1.25) has a unique solution, which is called the incenter of K and it is denoted by $\operatorname{inc}(K)$. General information concerning the theory of inradiuses and incenters of proper cones can be found in [6, 7].

The next corollary is an extension of [7, Theorem 3.3]. The symbol $\mathbf{1}_r$ refers to the r-dimensional vector of ones.

Corollary 1.12. *Let $(\mathbb{V}, \circ, \langle\cdot,\cdot\rangle)$ be a scalarizable EJA with rank r. Let Q be a permutation invariant proper cone in $\mathbb{R}^r$. Then*

(a) The spectral set $K = \{x \in \mathbb{V} : \lambda(x) \in Q\}$ is a proper cone in $\mathbb{V}$.

(b) K has the same inradius as Q, i.e.,

$$\rho(K) = \sup_{\substack{\mu \in Q \\ \|\mu\|_2=1}} \operatorname{dist}[\mu, \operatorname{bd}(Q)]. \tag{1.26}$$

(c) The incenter of K is given by

$$\operatorname{inc}(K) = \begin{cases} \frac{e}{\|e\|} & \text{if } \mathbf{1}_r \in Q \\ -\frac{e}{\|e\|} & \text{if } \mathbf{1}_r \notin Q. \end{cases} \tag{1.27}$$

Proof. The statement (a) is part of the folklore on spectral sets. Note that

$$\operatorname{bd}(K) = \{x \in \mathbb{V} : \lambda(x) \in \operatorname{bd}(Q)\}$$

is a spectral set because $\operatorname{bd}(Q)$ is permutation invariant. By using Proposition 1.11 one gets

$$\rho(K) = \sup_{\substack{x \in \mathbb{V},\ \lambda(x) \in Q \\ \sqrt{\theta}\,\|\lambda(x)\|_2=1}} \sqrt{\theta}\,\operatorname{dist}[\lambda(x), \operatorname{bd}(Q)],$$

where θ is the scaling factor of the EJA. The change of variables $\mu = \sqrt{\theta}\,\lambda(x)$ and the positive homogeneity of the function $\operatorname{dist}[\,\cdot\,, \operatorname{bd}(Q)]$ lead to (1.26). We now prove (c). Let $\bar{x} = \operatorname{inc}(K)$, that is to say,

$$\bar{x} \in K, \quad \|\bar{x}\| = 1, \quad \rho(K) = \operatorname{dist}[\bar{x}, \operatorname{bd}(K)].$$

Hence,

$$\lambda(\bar{x}) \in Q, \quad \sqrt{\theta}\|\lambda(\bar{x})\|_2 = 1, \quad \rho(Q) = \sqrt{\theta}\,\operatorname{dist}[\lambda(\bar{x}), \operatorname{bd}(Q)].$$

We have proven in this way that

$$\mathrm{inc}(Q) = \sqrt{\theta}\,\lambda(\bar{x}). \tag{1.28}$$

The proper cone Q being permutation invariant, one knows (cf.[7, Proposition 2.19]) that

$$\mathrm{inc}(Q) = \begin{cases} \frac{1}{\sqrt{r}}\,\mathbf{1}_r & \text{if} \quad \mathbf{1}_r \in Q \\ -\frac{1}{\sqrt{r}}\,\mathbf{1}_r & \text{if} \quad \mathbf{1}_r \notin Q. \end{cases} \tag{1.29}$$

There are two cases for consideration: if $\mathbf{1}_r \in Q$, then the combination of (1.28) and (1.29) yields

$$\lambda_1(\bar{x}) = \ldots = \lambda_r(\bar{x}) = \frac{1}{\sqrt{r\theta}}\,,$$

and therefore $\bar{x}$ is a positive multiple of e. If $\mathbf{1}_r \notin Q$, then

$$\lambda_1(\bar{x}) = \ldots = \lambda_r(\bar{x}) = -\frac{1}{\sqrt{r\theta}}\,,$$

and $\bar{x}$ is a negative multiple of e. This proves (1.27). □

Example 1.13. For instance, the cone of squares

$$K_{\mathbb{V}} \;:=\; \{x^2 : x \in \mathbb{V}\} \;=\; \{x \in \mathbb{V} : \lambda(x) \in \mathbb{R}^r_+\}$$

of an scalarizable EJA fits into the framework of Corollary 1.12. One gets

$$\rho(K_{\mathbb{V}}) = \frac{1}{\sqrt{r}} \quad \text{and} \quad \mathrm{inc}(K_{\mathbb{V}}) = \frac{e}{\|e\|}\,.$$

Bibliography

[1] F. Alizadeh and S.H. Schmieta. Extension of primal-dual interior point algorithms to symmetric cones. *Math. Program.*, 96 (2003), Ser. A, 409–438.

[2] J.P. Aubin and H. Frankowska. *Set-Valued Analysis*, Birhäuser, Boston, 1990.

[3] A. Daniilidis, A. Lewis, J. Malick, and H. Sendov. Prox-regularity of spectral functions and spectral sets. *J. Convex Anal.*, 15 (2008), 547–560.

[4] J. Faraut and A. Korányi. *Analysis on Symmetric Cones*. Clarendon Press, Oxford, 1994.

[5] M.S. Gowda, R. Sznajder, and J. Tao. Some P-properties for linear transformations on Euclidean Jordan algebras. *Linear Algebra Appl.*, 393 (2004), 203–232.

[6] R. Henrion and A. Seeger. On properties of different notions of centers for convex cones. *Set-Valued Var. Anal.*, 18 (2010), 205–231.

[7] R. Henrion and A. Seeger. Inradius and circumradius of various convex cones arising in applications. *Set-Valued Var. Anal.*, 18 (2010), 483–511.

[8] U. Hirzebruch. Über Jordan-Algebren und kompakte Riemannsche symmetrische Räume vom Rang 1. *Math. Z.*, 90 (1965), 339–354.

[9] A. Iusem and A. Seeger. Angular analysis of two classes of non-polyhedral convex cones: the point of view of optimization theory. *Comput. Appl. Math.*, 26 (2007), 191–214.

[10] A. Iusem and A. Seeger. Distances between closed convex cones: old and new results. *J. Convex Anal.*, 17 (2010), 1033–1055.

[11] A.S. Lewis. Convex analysis on the Hermitian matrices. *SIAM J. Optim.*, 6 (1996), 164–177.

[12] Y. Lim, J. Kim, and L. Faybusovich. Simultaneous diagonalization on simple Euclidean Jordan algebras and its applications. *Forum Math.* 15 (2003), 639–644.

[13] H. Ramírez, A. Seeger and D. Sossa. Commutation principle for variational problems on Euclidean Jordan algebras. *SIAM J. Optim.* 23 (2013), 687–694.

[14] A. Seeger. Convex analysis of spectrally defined matrix functions. *SIAM J. Optim.*, 7 (1997), 679–696.

Chapter 2

Complementarity problems with respect to Loewnerian cones[1]

Alberto Seeger[2] and David Sossa[3]

Abstract. This work deals with the analysis and numerical resolution of a broad class of complementarity problems on spaces of symmetric matrices. The complementarity conditions are expressed in terms of the Loewner ordering or, more generally, with respect to a dual pair of Loewnerian cones.

Mathematical subject classification: 15A18, 65F20, 65H10.

Key words: Nonlinear complementarity problem, Loewner ordering, cone-constrained eigenvalue problem, semismooth Newton method.

2.1 Introduction

The main concern of this work is the analysis and numerical resolution of a class of nonlinear complementarity problems formulated in $\mathbb{S}^n$, the space of symmetric matrices of order n. As

[1] The paper corresponding to this chapter was accepted for publication on *Journal of Global Optimization*, see [26].

[2]University of Avignon, Department of Mathematics, 33 rue Louis Pasteur, 84000 Avignon, France (E-mail: alberto.seeger@univ-avignon.fr).

[3]Departamento de Ingeniería Matemática, Centro de Modelamiento Matemático (CNRS UMI 2807), FCFM, Universidad de Chile, Blanco Encalada 2120, Santiago, Chile (E-mail: dsossa@dim.uchile.cl). This author is supported by CONICYT (Chile).

usual, $\mathbb{S}^n$ is equipped with the trace inner product $\langle Y, X\rangle = \mathrm{tr}(YX)$ and the associated norm. The first part of this work deals with a complementarity problem of the form

$$\text{(SDCP)} \quad \begin{cases} \Phi(X, Y, \lambda) = \mathbf{0} \\ \mathbf{0} \preceq X \perp Y \succeq \mathbf{0}, \end{cases} \tag{2.1}$$

where the nonnegativity constraints are expressed in terms of the Loewner ordering $\succeq$ on $\mathbb{S}^n$. Following [21], we refer to the equilibrium model (2.1) as the Semi-Definite Complementarity Problem (SDCP). The symbol λ stands for an unknown parameter vector in a Euclidean space Λ and Φ is a continuously differentiable function from $\mathbb{E} := \mathbb{S}^n \times \mathbb{S}^n \times \Lambda$ to another Euclidean space $\mathbb{F}$. The dimension of each bold marked zero vector $\mathbf{0}$ is understood from the context.

Although it is not strictly necessary, for simplicity in the exposition we assume that

$$\dim(\mathbb{F}) = \dim(\Lambda) + \dim(\mathbb{S}^n). \tag{2.2}$$

Such a dimensionality requirement is automatically satisfied in many practical cases. Below we display two concrete examples for which the assumption (2.2) is in force.

Example 2.1. Let $\mathrm{End}(\mathbb{S}^n)$ denote the vector space of linear endomorphisms on $\mathbb{S}^n$. A Loewner-eigenvalue of $\mathfrak{L} \in \mathrm{End}(\mathcal{S}_n)$ is a scalar $\lambda \in \mathbb{R}$ such that the system

$$\mathbf{0} \preceq X \perp (\mathfrak{L}(X) - \lambda X) \succeq \mathbf{0}$$

has a nonzero solution $X \in \mathbb{S}^n$. Finding a Loewner- eigenvalue of $\mathfrak{L}$ amounts to solve the complementarity problem

$$\begin{cases} \mathfrak{L}(X) - \lambda X - Y = \mathbf{0} \\ \mathrm{tr}(X) - 1 = 0 \\ \mathbf{0} \preceq X \perp Y \succeq \mathbf{0}. \end{cases} \tag{2.3}$$

The second equality in (2.3) is a normalization condition which ensures that X is a nonzero solution. In this example one has $\Lambda = \mathbb{R}$ and $\mathbb{F} = \mathbb{S}^n \times \mathbb{R}$. Theorem 2.1 in [25] ensures that the system (2.3) admits always a solution. The problem of finding Loewner-eigenvalues for some particular linear endomorphisms is addressed for instance in [31].

Example 2.2. Consider an optimization problem of the form

$$\text{minimize } c(X) \text{ s.t. } X \succeq \mathbf{0},\ \mathcal{A}(X) = b, \tag{2.4}$$

where $c : \mathbb{S}^n \to \mathbb{R}$ is a twice continuously differentiable convex function, $\mathcal{A} : \mathbb{S}^n \to \mathbb{R}^m$ is a linear map, and $b \in \mathbb{R}^m$. A particular instance of (2.4) is the so-called nearest correlation matrix problem, see [6, 17, 19]. The KKT- optimality conditions for the minimization problem (2.4) are

$$\begin{cases} \nabla c\,(X) - \mathcal{A}^T(\lambda) - Y = \mathbf{0} \\ \mathcal{A}(X) - b = \mathbf{0} \\ \mathbf{0} \preceq X \perp Y \succeq \mathbf{0}, \end{cases}$$

where $\nabla c : \mathbb{S}^n \to \mathbb{S}^n$ is the gradient map of c and $\mathcal{A}^T$ is the adjoint map of $\mathcal{A}$. In this example one has $\Lambda = \mathbb{R}^m$ and $\mathbb{F} = \mathbb{S}^n \times \mathbb{R}^m$.

Remark 2.3. Other particular cases of (2.1) can be found in [7, 13, 21]. Let $Q \in \mathbb{S}^n$ and $\mathfrak{L} \in \mathrm{End}(\mathbb{S}^n)$. The so-called semidefinite linear complementarity problem

$$\begin{cases} \mathfrak{L}(X) + Q - Y = \mathbf{0} \\ \mathbf{0} \preceq X \perp Y \succeq \mathbf{0} \end{cases}$$

also fits into the model (2.1). Semidefinite LCPs have been analyzed by numerous authors, both from a theoretical and algorithmic point of view.

The second part of our work deals with a complementarity problem having the more general form

$$\begin{cases} \Phi(X, Y, \lambda) = \mathbf{0} \\ \mathcal{K} \ni X \perp Y \in \mathcal{K}^*. \end{cases} \tag{2.5}$$

The nonnegativity constraints on X and Y are now expressed in terms a Loewnerian cone $\mathcal{K}$ and its positive dual cone $\mathcal{K}^*$. By definition, a Loewnerian cone is the image of the SDP cone

$$\mathbb{S}^n_+ := \{X \in \mathbb{S}^n : X \succeq \mathbf{0}\}$$

under some invertible linear endomorphism on $\mathbb{S}^n$. Loewnerian cones are not self-dual in general, but they share a number of properties of the SDP cone. For instance, a Loewnerian cone has a similar facial structure as the SDP cone.

2.2 Complementarity function approach for solving the SDCP

A natural strategy for solving (2.1) is to apply the Semismooth Newton Method (SNM) to the nonlinear system

$$\begin{cases} \Phi(X, Y, \lambda) = \mathbf{0} \\ \kappa(X, Y) = \mathbf{0}, \end{cases} \tag{2.6}$$

where $\kappa : \mathbb{S}^n \times \mathbb{S}^n \to \mathbb{S}^n$ is a complementarity function for the SDP cone, i.e.,

$$\kappa(X, Y) = \mathbf{0} \quad \Leftrightarrow \quad \mathbf{0} \preceq X \perp Y \succeq \mathbf{0}.$$

The function κ may not be differentiable. In order to apply the SNM one just needs to ensure that κ is globally Lipschitz and semismooth. This requirement is fulfilled if one chooses for instance the Fischer-Burmeister complementarity function

$$\kappa_{\mathrm{fb}}(X, Y) := X + Y - (X^2 + Y^2)^{1/2}$$

or the minimum complementarity function

$$\kappa_{\min}(X, Y) := X - \Pi_{\mathbb{S}^n_+}(X - Y)\,.$$

The square root operation $(\,\cdot\,)^{1/2}$ is understood in the usual matrix sense and $\Pi_{\mathbb{S}^n_+}$ stands for the projection map onto $\mathbb{S}^n_+$. The functions κ_{fb} and $\kappa_{\min}$ are known to be globally Lipschitz and strongly semismooth, see [28] and [29].

For notational convenience we write the nonlinear system (2.6) in the more compact form

$$\Psi(z) = \mathbf{0}, \tag{2.7}$$

where $\Psi : \mathbb{E} \to \mathbb{W}$ is given by

$$z = (X, Y, \lambda) \mapsto \Psi(z) := (\Phi(X, Y, \lambda), \kappa(X, Y)).$$

Thanks to the assumption (2.2), the space $\mathbb{W} := \mathbb{F} \times \mathbb{S}^n$ has the same dimension as $\mathbb{E}$. In other words, the system (2.7) has the same number of equations and unknown variables. We solve the square system (2.7) with the following Semismooth Newton Method (SNM):

- *Initialization.* Choose an initial point z_0 and set $t = 0$.
- *Iteration.* One has a current point z_t. Pick $M_t \in \partial\Psi(z_t)$, find Δz such that

$$M_t \Delta z = -\Psi(z_t), \tag{2.8}$$

  and update $z_{t+1} = z_t + \Delta z$.

The symbol $\partial\Psi(z)$ stands for the Clarke generalized Jacobian of Ψ at z (cf. [8]). Theoretical aspects concerning the rate of convergence of the SNM can be consulted in [24].

Remark 2.4. We keep the SNM running until any of the following situations occur:

$$\begin{array}{ll} t = 1000 & \text{(lack of convergence)}, \\ \text{cond}(M_t) \geq 10^{10} & \text{(ill-conditioning)}, \\ \|\Psi(z_t)\| \leq 10^{-8} & \text{(a solution is found)}. \end{array}$$

Here, $\text{cond}(M)$ refers to the condition number of a linear map M. All our numerical experiments are carried in Window 7 with a processor 3.40GHz Intel Xeon, Memory(RAM) 8.00 Gb. The codes were implemented with Matlab 7.12.

2.2.1 Numerical experiments

The numerical experiments reported in this section concern the problem of finding a Loewner-eigenvalue of a given map $\mathfrak{L} \in \text{End}(\mathbb{S}^n)$. So, the problem at hand is to solve (2.7) with $z = (X, Y, \lambda)$ and

$$\Psi(z) = (\mathfrak{L}(X) - \lambda X - Y, \text{tr}(X) - 1, \kappa(X, Y)). \tag{2.9}$$

The linearized system (2.8) takes here the particular form

$$\begin{aligned} \mathfrak{L}(\Delta X) - \lambda_t \Delta X - \Delta Y - (\Delta\lambda) X_t &= -[\mathfrak{L}(X_t) - \lambda_t X_t - Y_t] && (2.10) \\ \text{tr}(\Delta X) &= -[\text{tr}(X_t) - 1] && (2.11) \\ E_t(\Delta X) + F_t(\Delta Y) &= -\kappa(X_t, Y_t) && (2.12) \end{aligned}$$

with $(E_t, F_t) \in \partial\kappa(X_t, Y_t)$. A convenient way to initialize the SNM for finding a zero of the function (2.9) is to generate a random Gaussian matrix $\Xi \in \mathbb{S}^n$ and set:

$$\begin{aligned} X_0 &= [\mathrm{tr}(\Xi)]^{-1}\Xi, \\ \lambda_0 &= \|X_0\|^{-2}\,\langle \mathfrak{L}(X_0), X_0\rangle, \\ Y_0 &= \mathfrak{L}(X_0) - \lambda_0 X_0. \end{aligned}$$

That a random matrix $\Xi \in \mathbb{S}^n$ is Gaussian means that the entries $\Xi_{i,j}$ (with $i \leq j$) are random variables following a standard normal law.

Performance of κ_{fb}

As a first choice, we let κ be the Fisher-Burmeister complementarity function. For the reader's convenience, we record below a useful lemma that can be found in [18].

Lemma 2.5. *Let $A, B \in \mathbb{S}^n$ be such that $A^2 + B^2$ is nonsingular. Then κ_{fb} is continuously differentiable at (A, B) and the partial differentials (with respect to X and Y) are given by*

$$(D_X\kappa_{\mathrm{fb}})(A,B) = \mathfrak{I}_{\mathbb{S}^n} - \mathcal{L}^{-1}_{[A^2+B^2]^{1/2}} \circ \mathcal{L}_A \tag{2.13}$$

$$(D_Y\kappa_{\mathrm{fb}})(A,B) = \mathfrak{I}_{\mathbb{S}^n} - \mathcal{L}^{-1}_{[A^2+B^2]^{1/2}} \circ \mathcal{L}_B. \tag{2.14}$$

Here, $\mathfrak{I}_{\mathbb{S}^n}$ is the identity map on $\mathbb{S}^n$ and

$$U \in \mathbb{S}^n \mapsto \mathcal{L}_C(U) = C \bullet U := (CU + UC)/2$$

stands for the Lyapunov operator associated to $C \in \mathbb{S}^n$.

When κ is taken as the Fisher-Burmeister complementarity function, the equation (2.12) reads

$$E_t(\Delta X) + F_t(\Delta Y) = -\kappa_{\mathrm{fb}}(X_t, Y_t) \tag{2.15}$$

with $(E_t, F_t) \in \partial\kappa_{\mathrm{fb}}(X_t, Y_t)$. An explicit formula for computing the Clarke subdifferential of κ_{fb} can be found in [30]. If the matrix $X_t^2 + Y_t^2$ is nonsingular, then Lemma 2.5 yields

$$\begin{aligned} E_t &= \mathfrak{I}_{\mathbb{S}^n} - \mathcal{L}^{-1}_{C_t} \circ \mathcal{L}_{X_t}, \\ F_t &= \mathfrak{I}_{\mathbb{S}^n} - \mathcal{L}^{-1}_{C_t} \circ \mathcal{L}_{Y_t}, \end{aligned}$$

where $C_t := [X_t^2 + Y_t^2]^{1/2}$. In such a case, (2.15) takes the form

$$\Delta X - \mathcal{L}^{-1}_{C_t}(\mathcal{L}_{X_t}(\Delta X)) + \Delta Y - \mathcal{L}^{-1}_{C_t}(\mathcal{L}_{Y_t}(\Delta Y)) = -\kappa_{\mathrm{fb}}(X_t, Y_t).$$

By applying $\mathcal{L}_{C_t}$ on each side of the above equality and rearranging terms, one gets

$$(X_t - [X_t^2 + Y_t^2]^{1/2}) \bullet \Delta X + (Y_t - [X_t^2 + Y_t^2]^{1/2}) \bullet \Delta Y = [X_t^2 + Y_t^2]^{1/2} \bullet \kappa_{\mathrm{fb}}(X_t, Y_t). \tag{2.16}$$

In our first two experiments (cf. Tables 2.1 and 2.2), the map $\mathfrak{L}$ is generated by means of a random mechanism. To be more precise, we suppose that $\mathfrak{L}$ is Gaussian, i.e., the entries

of $\mathfrak{L}$ are independent random variables distributed according to a standard normal law. The "entries" of $\mathfrak{L}$ refer to the entries of the matrix representation of $\mathfrak{L}$ with respect to the canonical basis $\{E^{i,j}\}_{1\leq i\leq j\leq n}$ of $\mathbb{S}^n$. If $\{e_i\}_{i=1}^n$ denotes the canonical basis of $\mathbb{R}^n$, then

$$E^{i,j} := \begin{cases} e_i e_i^T & \text{if } i = j, \\ (e_i e_j^T + e_j e_i^T)/\sqrt{2} & \text{if } i \neq j. \end{cases}$$

The dimension of the space $\mathrm{End}(\mathbb{S}^n)$ is equal to t_n^2, where

$$t_n := \dim(\mathbb{S}^n) = n(n+1)/2.$$

So, one needs exactly t_n^2 scalars for defining $\mathfrak{L}$. The percentages reported in Table 2.1 are estimated by working with a sample of 10^4 Gaussian random maps $\mathfrak{L}$. Figures are rounded to one decimal place.

	$n=3$	$n=6$	$n=9$	$n=12$
success	96.7%	99.5%	99.2%	97.4%
divergence	2.3%	0.4%	0.5%	2.0%
ill-conditioning	1.0%	0.1%	0.3%	0.6%

Table 2.1: SNM applied to (2.9) with $\kappa = \kappa_{\mathrm{fb}}$. Percentages of success and failure.

The word "success" in Table 2.1 means that the selected random initial point z_0 has led to a solution, i.e., to a root of (2.9). Of course, as in any Newton type method, one can perfectly well encounter a situation of failure (divergence or ill-conditioning). As one can see from Table 2.1, the cases of ill-conditioning and divergence do not occur very often. We mention in passing that we never encountered a situation of nonsmoothness, i.e., at each visited point $z_t = (X_t, Y_t, \lambda_t)$, the matrix $X_t^2 + Y_t^2$ was always nonsingular.

In order to find a Loewner-eigenvalue of a given $\mathfrak{L}$, one initial point if often enough. On some occasions, however, one has to run the SNM with more than one initial point. Table 2.2 displays the expected Number of Initial Points (NIPs) needed for detecting a solution. These figures have been estimated by using a sample of 10^4 Gaussian random maps $\mathfrak{L}$, so they correspond to average values. The expected (or average) CPU time needed for detecting a solution increases of course with n.

	$n=3$	$n=6$	$n=9$	$n=12$
NIP	1.05	1.01	1.01	1.04
CPU	0.2	0.6	2.9	8.4

Table 2.2: SNM applied to (2.9) with $\kappa = \kappa_{\mathrm{fb}}$. Average NIPs and average CPU times needed for detecting a solution.

Remark 2.6. Note that (2.10)-(2.12) is a system of n^2+n+1 equations in the same number of unknown variables. So, the order of the system increases quadratically with n. This explains somehow the rapid growth in CPU time reported in Table 2.2.

The results reported in Table 2.1 are extremely encouraging. However, one should not be overoptimistic when $\mathfrak{L}$ belongs to some special subsets of the space $\mathrm{End}(\mathbb{S}^n)$. We have found that the percentages of success of the SNM are significatively lower if $\mathfrak{L}$ has for instance the special structure

$$\mathfrak{L}(X) = AXA + \langle C, X\rangle B, \tag{2.17}$$

where $A, B, C \in \mathbb{S}^n$.

Remark 2.7. By the way, as shown in the Appendix, it is possible to construct a triplet (A, B, C) for which the map (2.17) has infinitely many Loewner-eigenvalues. Such situation is however unlikely to occur if the matrices A, B, C are randomly generated.

Table 2.3 has been constructed in the same way as Table 2.1, except that now $\mathfrak{L}$ has the special form (2.17), where $A, B, C \in \mathbb{S}^n$ are random Gaussian matrices.

	$n = 3$	$n = 6$	$n = 9$	$n = 12$
success	89.9%	23.9%	4.3%	0.5%
divergence	6.6%	0.8%	0.1%	0.0%
ill-conditioning	3.5%	75.3%	95.6%	99.5%
nonsmoothness	3.5%	4.0%	2.4%	1.0%

Table 2.3: SNM applied to (2.9) with $\kappa = \kappa_{\mathrm{fb}}$. The map $\mathfrak{L}$ has the special form (2.17).

Two comments on Table 2.3 are in order.

i) Sometimes the current point $z_t = (X_t, Y_t, \lambda_t)$ is such that $X_t^2 + Y_t^2$ is singular. As shown in the last row of Table 2.3, such situation occurs but it is rather rare. In our opinion, it is not worthwhile to discuss in this paper the sophisticated issue of selecting a pair (E_t, F_t) in $\partial\kappa_{\mathrm{fb}}(X_t, Y_t)$. Instead of bothering with the computation of the set $\partial\kappa_{\mathrm{fb}}(X_t, Y_t)$, we simply use the equation (2.16), which makes sense even if $X_t^2 + Y_t^2$ is singular. The fact of encountering a point of nonsmoothness does not lead necessarily to a situation of failure.

ii) The instances of failure are essentially due to ill-conditioning, and not to divergence. This phenomenon is reinforced when n increases, see the last column of Table 2.3. Fixing the ill-conditioning problem induced by the special map (2.17) is not however the main concern of this work.

In spite of the negative results reported in Table 2.3, one can still consider the SNM as a viable strategy for computing the Loewner-eigenvalues of the map (2.17). Of course, one must accept the possibility of initializing the SNM with a large number of initial points. As shown in Table 2.4, the expected NIPs needed for detecting a solution grow rapidly with n.

	$n=3$	$n=6$	$n=9$	$n=12$
NIP	1.14	4.48	33.07	454.20
CPU	0.5	2.2	16.2	614.0

Table 2.4: SNM applied to (2.9) with $\kappa = \kappa_{\text{fb}}$. The map $\mathfrak{L}$ has the special form (2.17).

Performance of $\kappa_{\min}$

We now turn the attention to the minimum complementarity function. The equation (2.12) takes the form

$$E_t(\Delta X) + F_t(\Delta Y) = -\kappa_{\min}(X_t, Y_t) \tag{2.18}$$

with $(E_t, F_t) \in \partial\kappa_{\min}(X_t, Y_t)$. We shall need following differentiability lemma due to Malick and Sendov [20]. The notation $\mathbb{O}(n)$ stands for the set of orthogonal matrices of order n, $\lambda(X)$ refers to the n-dimensional vector whose components are the eigenvalues of $X \in \mathbb{S}^n$ arranged in nondecreasing order, $\text{Diag}(x)$ is the diagonal matrix whose diagonal entries are the components of $x \in \mathbb{R}^n$, and $\odot$ is used to indicate the Hadamard (or componentwise) product. For each $x \in \mathbb{R}^n$ such that $\Pi_{i=1}^n x_i \neq 0$ and $x_1 \leq \ldots \leq x_n$, one defines $\mathcal{B}(x) \in \mathbb{S}^n$ by setting

$$(\mathcal{B}(x))_{i,j} := \begin{cases} 0 & \text{if } i \leq q, j \leq q, \\ 1 & \text{if } i > q, j > q, \\ x_j/(x_j - x_i) & \text{if } i \leq q, j > q, \\ x_i/(x_i - x_j) & \text{if } i > q, j \leq q, \end{cases}$$

where q is the number of negative components of x.

Lemma 2.8 (Malick and Sendov, 2006)**.** *The function $\mathcal{M} : \mathbb{S}^n \to \mathbb{R}$, defined by*

$$\mathcal{M}(X) := \min_{Y \preceq \mathbf{0}} \frac{1}{2}\|X - Y\|^2,$$

is differentiable and its gradient at X is equal to $\Pi_{\mathbb{S}^n_+}(X)$. Furthermore, $\mathcal{M}$ is twice differentiable at X if and only if X is nonsingular, in which case

$$D^2\mathcal{M}(X)(H_1, H_2) = \langle \mathcal{B}(\lambda(X)), (U^T H_1 U) \odot (U^T H_2 U) \rangle, \tag{2.19}$$

where $U \in \mathbb{O}(n)$ is such that $X = U\text{Diag}(\lambda(X))U^T$.

Thus, for all nonsingular $X \in \mathbb{S}^n$, one has

$$D^2\mathcal{M}(X)(H_1, H_2) = \langle D\Pi_{\mathbb{S}^n_+}(X)(H_1), H_2 \rangle$$

and computing $D\Pi_{\mathbb{S}^n_+}(X)(H)$ is nothing but to find the Riesz representation of the linear functional $D^2\mathcal{M}(X)(H, \cdot)$. By using the formula (2.19), one gets

$$\begin{aligned} D\Pi_{\mathbb{S}^n_+}(X)(H) &= U\left[\mathcal{B}(\lambda(X)) \odot (U^T H U)\right] U^T \\ &= \sum_{1 \leq i,j \leq n} \langle \mathcal{B}(\lambda(X)), (U^T H U) \odot (U^T E^{i,j} U) \rangle E^{i,j}. \end{aligned} \tag{2.20}$$

For numerical purposes, we rely on the representation (2.20).

Lemma 2.9. *Let $A, B \in \mathbb{S}^n$ be such that $A - B$ is nonsingular. The $\kappa_{\min}$ is differentiable at (A, B) and the partial differentials (with respect to X and Y) are given by*

$$\begin{aligned}(D_X\kappa_{\min})(A,B) &= \mathfrak{I}_{\mathbb{S}^n} - D\Pi_{\mathbb{S}^n_+}(A-B),\\ (D_Y\kappa_{\min})(A,B) &= D\Pi_{\mathbb{S}^n_+}(A-B).\end{aligned}$$

By exploiting the formulas established in Lemma 2.9, one sees that (2.18) takes the form

$$[\mathfrak{I}_{\mathbb{S}^n} - D\Pi_{\mathbb{S}^n_+}(X_t - Y_t)](\Delta X) + D\Pi_{\mathbb{S}^n_+}(X_t - Y_t)(\Delta Y) = -\kappa_{\min}(X_t, Y_t).$$

whenever $X_t - Y_t$ is nonsingular. Tables 2.5 and 2.6 have been constructed in the same way as Tables 2.1 and 2.2, respectively, except that now κ is the minimum complementarity function. The map $\mathfrak{L}$ is Gaussianly generated and has no special structure.

	$n = 3$	$n = 6$	$n = 9$	$n = 12$
success	86.7%	97.6%	99.3%	98.6%
divergence	12.9%	2.4%	0.6%	1.0%
ill-conditioning	0.4%	0.0%	0.1%	0.4%

Table 2.5: SNM applied to (2.9) with $\kappa = \kappa_{\min}$.

	$n = 3$	$n = 6$	$n = 9$	$n = 12$
NIP	1.23	1.04	1.01	1.04
CPU	0.9	2.6	13.7	67.0

Table 2.6: SNM applied to (2.9) with $\kappa = \kappa_{\min}$.

As in Table 2.1, we never encountered a situation of nonsmoothness. Table 2.5 shows that, in terms of percentages of success, the performance of $\kappa = \kappa_{\min}$ is of the same order as $\kappa = \kappa_{\text{fb}}$. By contrast, Table 2.6 shows that the CPU time needed by $\kappa = \kappa_{\min}$ is clearly higher than the CPU time needed by $\kappa = \kappa_{\text{fb}}$. The reason is that the computational effort for determining (E_t, F_t) in (2.18) is much higher than in (2.15).

Remark 2.10. The performance of $\kappa = \kappa_{\min}$ is similar to that of $\kappa = \kappa_{\text{fb}}$, also when $\mathfrak{L}$ has the special structure (2.17). For avoiding repetitions, we omit writing down the corresponding table.

2.2.2 A brief comment on the squaring technique

The squaring technique is based on the representability of the SDP cone as a cone of squares: $\mathbb{S}^n_+ = \{U^2 : U \in \mathbb{S}^n\}$. One can get rid of the constraints $X \succeq 0$ and $Y \succeq 0$ by writing $X = U^2$ and $Y = V^2$. With such change of variables, the model (2.1) takes the form of a smooth system of equations:

$$\begin{cases} \Phi\left(U^2, V^2, \lambda\right) = \mathbf{0}, \\ \langle U^2, V^2 \rangle = 0. \end{cases} \tag{2.21}$$

Unfortunately, the last equality in (2.21) is at the origin of a certain ill-conditioning in the whole system. Besides, the system (2.21) is not square. By following a similar strategy as in [9, 10], one may shift the attention to a certain "companion" system

$$\begin{cases} \Phi(U^2, V^2, \lambda) = \mathbf{0}, \\ U \bullet V = \mathbf{0}, \end{cases} \tag{2.22}$$

which is smooth, square, and usually well-conditioned. It must be observed that (2.22) is related, but not equivalent, to the original problem (2.21). More precisely, $\langle U^2, V^2 \rangle = 0$ implies $U \bullet V = \mathbf{0}$, but not conversely. The triplet (U, V, λ) is declared a *fake solution* to (2.21) if the system (2.22) holds, but $\langle U^2, V^2 \rangle \neq 0$.

Example 2.11. Consider the problem of finding a Loewner-eigenvalue of the map $\mathcal{L}$ given by $\mathcal{L}(X) = \mathrm{tr}(X) I_n$, where I_n is the identity matrix of order n. The companion system (2.22) becomes

$$\begin{cases} I_n - \lambda U^2 - V^2 = \mathbf{0}, \\ \|U\|^2 - 1 = 0, \\ U \bullet V = \mathbf{0}. \end{cases} \tag{2.23}$$

One can check that

$$(U, V, \lambda) = \left(\frac{-e_1 e_1^T + e_n e_n^T}{\sqrt{2}}, \sum_{i=1}^{n} e_i e_{n+1-i}^T, 0 \right) \tag{2.24}$$

is a fake solution. Indeed, (2.24) solves the system (2.23), but

$$\langle U^2, V^2 \rangle = \left\langle \frac{e_1 e_1^T + e_n e_n^T}{2}, I_n \right\rangle = 1.$$

If one considers a map $\mathfrak{L} \in \mathrm{End}(\mathbb{S}^n)$ that is randomly generated according to a Gaussian distribution and applies the classical Newton method to the companion system

$$\begin{cases} \mathfrak{L}(U^2) - \lambda U^2 - V^2 = \mathbf{0}, \\ \|U\|^2 - 1 = 0, \\ U \bullet V = \mathbf{0}, \end{cases} \tag{2.25}$$

then one observes experimentally that, in case of convergence, one always obtains a triplet (U, V, λ) that is not a fake solution, but a true one. In other words, the delivery of a fake solution is a rather exceptional event. Instances of ill-conditioning in (2.25) can be observed from time to time, but not frequently. On the negative side, our numerical tests show that Newton's method applied to (2.25) requires a very careful selection of the initial point in order to ensure convergence. In view of this observation, it is reasonable to introduce a suitable globalization technique as recommended by some authors.

2.3 On Loewnerian cones

In what follows, $\mathbb{GL}(\mathbb{S}^n)$ stands for the group of invertible linear endomorphisms on $\mathbb{S}^n$, i.e.,

$$\mathbb{GL}(\mathbb{S}^n) := \{\mathfrak{F} \in \mathrm{End}(\mathbb{S}^n) : \mathfrak{F} \text{ is invertible}\}.$$

The next definition concerns a class of closed convex cones that are somewhat similar to $\mathbb{S}^n_+$.

Definition 2.12. *A closed convex cone $\mathcal{K}$ in $\mathbb{S}^n$ is* Loewnerian *if it representable as*

$$\mathcal{K} = \{\mathfrak{F}(U) : U \succeq \mathbf{0}\} \tag{2.26}$$

for some $\mathfrak{F} \in \mathbb{GL}(\mathbb{S}^n)$. One refers to (2.26) as the Loewnerian cone induced by $\mathfrak{F}$.

A Loewnerian cone is nothing but the image of $\mathbb{S}^n_+$ under some invertible linear endomorphism on $\mathbb{S}^n$. The map $\mathfrak{F}$ in the representation formula (2.26) is not unique. In fact,

$$\left.\begin{array}{l}\mathfrak{F}_1, \mathfrak{F}_2 \in \mathbb{GL}(\mathbb{S}^n) \text{ induce} \\ \text{the same Loewnerian cone}\end{array}\right\} \;\Leftrightarrow\; (\mathfrak{F}_2^{-1} \circ \mathfrak{F}_1)(\mathbb{S}^n_+) = \mathbb{S}^n_+.$$

Since $\mathbb{S}^n_+$ is proper and nonpolyhedral, so is any Loewnerian cone. A closed convex cone is called proper if it is pointed and solid. A Loewnerian cone can also be represented in the "inverse image" form

$$\mathcal{K} = \{X \in \mathbb{S}^n : \mathfrak{G}(X) \succeq \mathbf{0}\} \tag{2.27}$$

for some $\mathfrak{G} \in \mathbb{GL}(\mathbb{S}^n)$. One passes from (2.26) to the dual representation (2.27) by taking $\mathfrak{G}$ as the inverse of $\mathfrak{F}$. Conversely, one passes from (2.27) to the primal representation (2.26) by taking $\mathfrak{F}$ as the inverse of $\mathfrak{G}$. The dual of a Loewnerian cone is a Loewnerian cone. Indeed, for all $\mathfrak{F} \in \mathbb{GL}(\mathbb{S}^n)$, one has

$$\{\mathfrak{F}(U) : U \succeq \mathbf{0}\}^* = \{\mathfrak{F}^{-T}(V) : V \succeq \mathbf{0}\} \tag{2.28}$$

with $\mathfrak{F}^{-T}$ standing for the adjoint map of the inverse of $\mathfrak{F}$. The following proposition concerns the facial structure of Loewnerian cones.

Proposition 2.13. *Let $\mathcal{K}$ be a Loewnerian cone induced by $\mathfrak{F} \in \mathbb{GL}(\mathbb{S}^n)$.*

(a) Each face of $\mathcal{K}$ is an exposed face. Furthermore, $\{\dim(\mathcal{M}) : \mathcal{M}$ face of $\mathcal{K}\} = \{t_1, \ldots, t_n\}$, where $t_k := k(k+1)/2$ stands for the k-th triangular number.

(b) $\mathcal{M}$ is a t_k-dimensional face of $\mathcal{K}$ if and only if $\mathcal{M} = \{\mathfrak{F}(U) : U \succeq \mathbf{0}, \operatorname{Im} U \subseteq L\}$ for some linear subspace $L \subseteq \mathbb{R}^n$ of dimension k.

Proof. As an application of [4, Theorem 5], one sees that the transformation

$$\begin{array}{rcl} 2^{\mathbb{S}^n} & \xrightarrow{\mathfrak{F}_\sharp} & 2^{\mathbb{S}^n} \\ \mathfrak{F}_\sharp(\mathcal{E}) & := & \{\mathfrak{F}(U) : U \in \mathcal{E}\} \end{array}$$

is a bijection between the faces of $\mathbb{S}^n_+$ and the faces of $\mathcal{K}$. One can also check that

$$\dim[\mathfrak{F}_\sharp(\mathcal{E})] = \dim(\mathcal{E})$$

for any face $\mathcal{E}$ of $\mathbb{S}^n_+$. The rest of the proof is a matter of recalling the well known facial structure of $\mathbb{S}^n_+$, see for instance [5, Chapter II.12] or [12, Section 4.2.2]. □

A proper cone $\mathcal{K}$ in $\mathbb{S}^n$ is called rotund (cf. [27]) if every face of $\mathcal{K}$, other than $\mathcal{K}$ itself and $\{\mathbf{0}\}$, is a half-line. Rotund cones are often times referred to as strictly convex cones because they are characterized by the strict convexity condition

$$X_1, X_2 \in \mathcal{K} \text{ not collinear} \implies X_1 + X_2 \in \operatorname{int}(\mathcal{K}). \tag{2.29}$$

By definition, a proper cone $\mathcal{K}$ in $\mathbb{S}^n$ is smooth if its dual is rotund.

Corollary 2.14. *A Loewnerian cone in $\mathbb{S}^n$, with $n \geq 3$, is neither rotund nor smooth.*

Proof. Suppose that $n \geq 3$. A rotund cone in $\mathbb{S}^n$ does not have a face of dimension $t_2 = 3$. Hence, it cannot be Loewnerian by Proposition 2.13(a). A smooth cone in $\mathbb{S}^n$ cannot be Loewnerian either, because its dual is rotund. □

2.3.1 Complementarity relative to Loewnerian cones

Consider the general complementarity problem (2.5) with $\mathcal{K}$ being a Loewnerian cone induced by $\mathfrak{F}$. Thanks to the duality formula (2.28) and the fact that

$$\langle \mathfrak{F}(U), \mathfrak{F}^{-T}(V) \rangle = \langle U, V \rangle$$

for all $U, V \in \mathbb{S}^n$, the model (2.5) can be written in the equivalent form

$$\begin{cases} \Phi(\mathfrak{F}(U), \mathfrak{F}^{-T}(V), \lambda) = \mathbf{0}, \\ \mathbf{0} \preceq U \perp V \succeq \mathbf{0}. \end{cases} \tag{2.30}$$

The latter complementarity problem has of course the same structure as (2.1). We are then back to the context of Section 2.2. One can solve (2.30) by using for instance the complementarity function technique with the Fisher-Burmeister function κ_{fb}. Another option is to use the complementarity function technique directly on (2.5). As complementarity function for the Loewnerian cone $\mathcal{K}$ one can use for instance

$$\widehat{\kappa}_{\mathrm{fb}}(X, Y) := \kappa_{\mathrm{fb}}\left(\mathfrak{F}^{-1}(X), \mathfrak{F}^{T}(Y)\right).$$

Example 2.15. Consider the problem of finding a real λ such that the system

$$\mathcal{K} \ni X \perp (\mathfrak{L}(X) - \lambda X) \in \mathcal{K}^*$$

has a nonzero solution $X \in \mathbb{S}^n$. If $\mathcal{K}$ is a Loewnerian cone induced by $\mathfrak{F}$, then everything boils down to solve the nonlinear system

$$\begin{cases} \mathfrak{L}(\mathfrak{F}(U)) - \lambda \mathfrak{F}(U) - \mathfrak{F}^{-T}(V) = \mathbf{0}, \\ \operatorname{tr}(\mathfrak{F}(U)) - 1 = 0, \\ \kappa_{\mathrm{bf}}(U, V) = \mathbf{0}. \end{cases}$$

Alternatively, one can work with the original variables X and Y:

$$\begin{cases} \mathfrak{L}(X) - \lambda X - Y = \mathbf{0}, \\ \operatorname{tr}(X) - 1 = 0, \\ \widehat{\kappa}_{\mathrm{fb}}(X, Y) = \mathbf{0}. \end{cases}$$

2.3.2 Examples of Loewnerian cones and counter-examples

We now address a short list of interesting proper cones that are Loewnerian. We do not claim that all these cones are relevant in the theory of complementarity problems, but it is reasonable to get acquainted with these examples just for the sake of academic knowledge. We start by stating a trivial but useful lemma.

Lemma 2.16. *Let $C, B \in \mathbb{S}^n$ be such that $\langle C, B\rangle \neq 1$. Then the map $\mathfrak{G} : \mathbb{S}^n \to \mathbb{S}^n$ defined by*

$$\mathfrak{G}(X) = \langle C, X\rangle\, B - X \tag{2.31}$$

is invertible and its inverse $\mathfrak{F} : \mathbb{S}^n \to \mathbb{S}^n$ is given by

$$\mathfrak{F}(U) = \frac{\langle C, U\rangle}{\langle C, B\rangle - 1}\, B - U.$$

Proof. For all $U \in \mathbb{S}^n$, the matrix equation $\langle C, X\rangle\, B - X = U$ admits a unique solution $X \in \mathbb{S}^n$. Indeed, by taking the inner product with respect to C one gets

$$\langle C, X\rangle\, (\langle C, B\rangle - 1) = \langle C, U\rangle.$$

From here one derives $\langle C, X\rangle$ as a function of U and then one finds the solution X itself. □

If X, B are symmetric matrices with B positive definite, then $\lambda_{\max}(X, B)$ denotes the largest real λ for which $\det(X - \lambda B) = 0$. When B is the identity matrix one simply writes $\lambda_{\max}(X)$.

Proposition 2.17. *Let $C, B \in \mathbb{S}^n$ with B positive definite. The condition $\langle C, B\rangle \neq 1$ is necessary and sufficient for the closed convex cone*

$$\mathcal{K} = \{X \in \mathbb{S}^n : \lambda_{\max}(X, B) \leq \langle C, X\rangle\} \tag{2.32}$$

to be Loewnerian.

Proof. Let $\mathfrak{G} : \mathbb{S}^n \to \mathbb{S}^n$ be given by (2.31). One has

$$\begin{aligned} \mathfrak{G}(X) \succeq \mathbf{0} &\Leftrightarrow \langle C, X\rangle B \succeq X \\ &\Leftrightarrow \langle u, Xu\rangle \leq \langle C, X\rangle\, \langle u, Bu\rangle \ \text{ for all } u \in \mathbb{R}^n \\ &\Leftrightarrow \lambda_{\max}(X, B) \leq \langle C, X\rangle. \end{aligned}$$

In other words, $\mathcal{K}$ is representable as in (2.27). If $\langle C, B\rangle \neq 1$, then $\mathfrak{G}$ is invertible by Lemma 2.16, and therefore $\mathcal{K}$ is Loewnerian. If $\langle C, B\rangle = 1$, then $\mathbb{R}(B) = \mathrm{Ker}(\mathfrak{G}) \subseteq \mathcal{K}$. Hence, $\mathcal{K}$ is not pointed because it contains the line generated by B. □

Remark 2.18. Suppose that $\langle C, B\rangle \neq 1$. The Loewnerian cone (2.32) is not self-dual in general. However, its dual

$$\mathcal{K}^* = \left\{Y \in \mathbb{S}^n : \lambda_{\max}(Y, C) \leq \frac{\langle B, Y\rangle}{\langle C, B\rangle - 1}\right\}$$

has the same structure as (2.32).

Corollary 2.19. *Let $\alpha \in \mathbb{R}\backslash\{0,1\}$. Then*

$$\begin{aligned} K &= \left\{X \in \mathbb{S}^n : \alpha\, \lambda_{\max}(X) \leq n^{-1}\mathrm{tr}(X)\right\} \\ K^* &= \left\{Y \in \mathbb{S}^n : (1-\alpha)\lambda_{\max}(Y) \leq n^{-1}\mathrm{tr}(Y)\right\} \end{aligned}$$

are mutually dual Loewnerian cones.

Proof. One just needs to apply Proposition 2.17 with $B = I_n$ and $C = (\alpha n)^{-1} I_n$. □

Recall that $\lambda(X) = (\lambda_1(X), \ldots, \lambda_n(X))^T$ is the vector whose components are the eigenvalues of $X \in \mathbb{S}^n$ arranged in nondecreasing order. A spectral proper cone in $\mathbb{S}^n$ is a set of the form

$$\mathcal{K} = \{X \in \mathbb{S}^n : \lambda(X) \in P\}, \tag{2.33}$$

where P is a permutation invariant proper cone in $\mathbb{R}^n$. An interesting family $\{\mathcal{K}_q^{\uparrow}\}_{q=1}^n$ of spectral proper cones arising in applications is given by

$$\mathcal{K}_q^{\uparrow} := \left\{ X \in \mathbb{S}^n : \sum_{i=1}^{q} \lambda_i(X) \geq 0 \right\}.$$

These cones have been studied in [3] and in many other publications. For $q = 1$, one gets

$$\mathcal{K}_1^{\uparrow} \;=\; \{X \in \mathbb{S}^n : \lambda_{\min}(X) \geq 0\} \;=\; \mathbb{S}^n_+.$$

The choice $q = n-1$ leads to

$$\mathcal{K}_{n-1}^{\uparrow} = \{X \in \mathbb{S}^n : \lambda_{\max}(X) \leq \mathrm{tr}(X)\},$$

which is also a Loewnerian cone (cf. Corollary 2.19). The analysis of the case $2 \leq q \leq n-2$ is more involved. The next proposition concerns the choice $q = 2$.

Proposition 2.20. *Let $n \geq 4$. The cone*

$$\mathcal{K}_2^{\uparrow} = \{X \in \mathbb{S}^n : \lambda_1(X) + \lambda_2(X) \geq 0\} \tag{2.34}$$

is not Loewnerian.

Proof. The cone $\mathcal{K}_2^{\uparrow}$ admits the representation (2.33) with

$$P = \{x \in \mathbb{R}^n : x_i + x_j \geq 0 \text{ for all } 1 \leq i < j \leq n\}.$$

Such set P is a permutation invariant proper cone. Note that P is also polyhedral. In view of [14, Lemma 2.1], the set

$$E = \{x \in P : x_i + x_{n-1} = 0 \text{ for all } i = 1, \ldots, n-2\}$$

is a two-dimensional exposed face of P. By using such face E and Lewis's facial theorem [16, Theorem 5.1], one can construct a two-dimensional face for $\mathcal{K}_2^{\uparrow}$. This fact and Proposition 2.13(a) prove that $\mathcal{K}_2^{\uparrow}$ is not Loewnerian. □

The cone $\mathcal{K}^{\uparrow}_{n-2}$ is not Loewnerian either, because it is the image of $\mathcal{K}^{\uparrow}_{2}$ under a nonsingular transformation. In fact, one has the following general result.

Proposition 2.21. *Let $q \in \{1, \ldots, n-1\}$. Then the linear map*

$$X \in \mathbb{S}^n \mapsto \mathfrak{G}(X) := \frac{\mathrm{tr}(X)}{q}\, I_n - X$$

is a bijection between $\mathcal{K}^{\uparrow}_{n-q}$ and $\mathcal{K}^{\uparrow}_{q}$.

Proof. That $\mathfrak{G}$ is invertible is clear from Lemma 2.16. Since

$$\lambda_i(\mathfrak{G}(X)) = \frac{\mathrm{tr}(X)}{q} - \lambda_{n-i+1}(X)$$

for all $i \in \{1, \ldots, n\}$, one has

$$\sum_{i=1}^{q} \lambda_i(\mathfrak{G}(X)) = \mathrm{tr}(X) - \sum_{i=1}^{q} \lambda_{n-i+1}(X) \;=\; \sum_{i=1}^{n-q} \lambda_i(X).$$

Hence,

$$\begin{aligned} \mathcal{K}^{\uparrow}_{n-q} &= \left\{ X \in \mathbb{S}^n : \sum_{i=1}^{n-q} \lambda_i(X) \geq 0 \right\} \\ &= \left\{ X \in \mathbb{S}^n : \sum_{i=1}^{q} \lambda_i(\mathfrak{G}(X)) \geq 0 \right\} = \mathfrak{G}^{-1}(\mathcal{K}^{\uparrow}_{q}). \end{aligned}$$

This completes the proof. □

By a revolution cone in $\mathbb{S}^n$ one understands a proper cone of the form

$$\mathrm{Rev}(s, Y) := \{U \in \mathbb{S}^n : s\,\|U\| \leq \langle Y, U\rangle\},$$

where $0 < s < 1$ and $Y \in \mathbb{S}^n$ is such that $\|Y\| = 1$. A revolution cone is a particular instance of an ellipsoidal cone. By definition, an ellipsoidal cone in $\mathbb{S}^n$ is a set representable as

$$\mathcal{E}(\mathfrak{D}, B) := \{X \in \mathbb{S}^n : \|\mathfrak{D}(X)\| \leq \langle B, X\rangle\}, \tag{2.35}$$

where $B \in \mathbb{S}^n$ and $\mathfrak{D} \in \mathbb{GL}(\mathbb{S}^n)$ are such that $\|\mathfrak{D}^{-T}(B)\| > 1$. The latter inequality ensures that (2.35) is a proper cone. Whether an ellipsoidal cone in $\mathbb{S}^n$ is Loewnerian or not depends on the dimension n. The case $n = 2$ is somewhat special.

Proposition 2.22. *One has:*

(a) A proper cone in $\mathbb{S}^2$ is Loewnerian if and only if it is an ellipsoidal cone.

(b) Any ellipsoidal cone in $\mathbb{S}^n$, with $n \geq 3$, is not Loewnerian.

Proof. One easily sees that (2.35) satisfies the strict convexity condition (2.29). Hence, (b) is a consequence of Corollary 2.14. Consider now the particular case $n = 2$. Note that

$$\mathcal{E}(\mathfrak{D}, B) = \left\{ \mathfrak{D}^{-1}(U) : U \in \mathrm{Rev}\left(\frac{1}{\|\mathfrak{D}^{-T}B\|}, \frac{\mathfrak{D}^{-T}(B)}{\|\mathfrak{D}^{-T}B\|} \right) \right\},$$

i.e., the action of $\mathfrak{D}$ on the ellipsoidal cone (2.35) produces a revolution cone. So, it is enough to show that a suitable chain

$$\mathbb{S}^2 \xrightarrow{\mathfrak{L}_1} \mathbb{R}^3 \xrightarrow{\mathfrak{L}_2} \mathbb{R}^3 \xrightarrow{\mathfrak{L}_3} \mathbb{S}^2$$

of linear isomorphisms allows us to pass from a revolution cone to $\mathbb{S}^2_+$. One starts with the standard linear isometry

$$\mathfrak{L}_1 \left(\begin{bmatrix} \alpha & \beta \\ \beta & \gamma \end{bmatrix} \right) := \begin{bmatrix} \alpha \\ \sqrt{2}\beta \\ \gamma \end{bmatrix}$$

between $\mathbb{S}^2$ and $\mathbb{R}^3$. The action of $\mathfrak{L}_1$ on a revolution cone $\mathrm{Rev}(s, Y)$ in $\mathbb{S}^2$ produces a revolution cone

$$\mathrm{rev}(s, y) := \{u \in \mathbb{R}^3 : s\,\|u\| \le y^T u\}$$

in $\mathbb{R}^3$. Then one constructs a linear invertible map $\mathfrak{L}_2 : \mathbb{R}^3 \to \mathbb{R}^3$ that converts $\mathrm{rev}(s, y)$ into the ice-cream cone

$$\mathbb{L}_3 := \{x \in \mathbb{R}^3 : [x_1^2 + x_2^2]^{1/2} \le x_3\}.$$

Such an $\mathfrak{L}_2$ clearly exists. Finally, one constructs a linear isomorphism $\mathfrak{L}_3 : \mathbb{R}^3 \to \mathbb{S}^2$ that converts $\mathbb{L}_3$ into $\mathbb{S}^2_+$. The explicit form of $\mathfrak{L}_3$ can be found in [12, Section 2.6] or in [22, Chapter 2.5]. □

The next proposition concerns two classes of closed convex cones arising in the stability analysis of dynamical systems.

Proposition 2.23. *Consider a matrix $A \in \mathbb{M}_n$ with possible complex eigenvalues $\mu_1, \ldots, \mu_n$. Then*

$$\mathcal{K}_{\mathrm{lyap}} := \{X \in \mathbb{S}^n : AX + XA^T \succeq \mathbf{0}\}, \tag{2.36}$$
$$\mathcal{K}_{\mathrm{stein}} := \{X \in \mathbb{S}^n : X \succeq A^T X A\} \tag{2.37}$$

are Loewnerian cones under the respective assumptions

$$\mu_i + \mu_j \neq 0 \quad \text{for all } i, j \in \{1, \ldots, n\}, \tag{2.38}$$
$$\mu_i\,\mu_j \neq 1 \quad \text{for all } i, j \in \{1, \ldots, n\}. \tag{2.39}$$

Proof. The cones (2.36) and (2.37) are representable as in (2.27) with $\mathfrak{G}$ given respectively by

$$\mathfrak{G}_{\mathrm{lyap}}(X) := AX + XA^T,$$
$$\mathfrak{G}_{\mathrm{stein}}(X) := X - A^T X A.$$

As established in [15], the spectral condition (2.38) implies the invertibility of the Lyapunov operator $\mathfrak{G}_{\mathrm{lyap}} : \mathbb{S}^n \to \mathbb{S}^n$, whereas (2.39) implies the invertibility of the Stein operator $\mathfrak{G}_{\mathrm{stein}} : \mathbb{S}^n \to \mathbb{S}^n$. □

2.4 By way of application: Finding the nearest Euclidean distance matrix

A square matrix A of order m is an Euclidean distance matrix (EDM) if there are points $\{p_i\}_{i=1}^m$ in some Euclidean space $\mathbb{R}^d$ such that

$$a_{ij} = \|p_i - p_j\|_2^2 \quad \text{for all } i, j \in \{1, \ldots, m\},$$

where $\|\cdot\|_2$ is the usual Euclidean norm on $\mathbb{R}^d$. The EDMs of order m form a proper cone in the space

$$\mathbb{S}_\bullet^m := \{X \in \mathbb{S}^m : \operatorname{diag}(X) = \mathbf{0}\},$$

where $\operatorname{diag}(X)$ denotes the vector whose components are the diagonal entries of X. There is a rich literature devoted to the analysis and applications of the cone

$$\mathcal{E}_m := \{A \in \mathbb{S}_\bullet^m : A \text{ is an EDM}\}.$$

The following proposition exploits the fact that $\mathbb{S}_\bullet^{n+1}$ has the same dimension as $\mathbb{S}^n$.

Proposition 2.24. *Let ϕ be a linear isomorphism between $\mathbb{S}_\bullet^{n+1}$ and $\mathbb{S}^n$. Then*

$$\mathcal{K} = \{\phi(A) : A \in \mathcal{E}_{n+1}\} \tag{2.40}$$

is a Loewnerian cone.

Proof. There exists a linear isomorphism $\mathfrak{L} : \mathbb{S}^n \to \mathbb{S}_\bullet^{n+1}$ such that

$$\mathcal{E}_{n+1} = \{\mathfrak{L}(U) : U \succeq \mathbf{0}\}.$$

As shown in [1, Theorem 3.2], one may consider for instance

$$\mathfrak{L}(U) := \begin{bmatrix} 0 & [\operatorname{diag}(U)]^T \\ \operatorname{diag}(U) & \mathfrak{B}(U) - 2U \end{bmatrix},$$

where

$$\mathfrak{B}(U) := [\operatorname{diag}(U)]\mathbf{1}_n^T + \mathbf{1}_n[\operatorname{diag}(U)]^T$$

and $\mathbf{1}_n$ is the n-dimensional vector of ones. Hence, (2.40) is a Loewnerian cone induced by the composition $\mathfrak{F} = \phi \circ \mathfrak{L}$. □

Remark 2.25. The dual of (2.40) is the Loewnerian cone induced by $\phi^{-T} \circ \mathfrak{L}^{-T}$. A matter of computation shows that $\mathfrak{L}^{-T} : \mathbb{S}^n \to \mathbb{S}_\bullet^{n+1}$ is given by

$$\mathfrak{L}^{-T}(V) = \frac{1}{2}\begin{bmatrix} 0 & (V\mathbf{1}_n)^T \\ V\mathbf{1}_n & -V^\diamond \end{bmatrix},$$

where $V^\diamond$ is derived from V by setting zero to the diagonal entries.

An interesting problem of numerical linear algebra addressed in [1, 11, 23] is this: Given a matrix $C \in \mathbb{S}_\bullet^{n+1}$ generated by a random mechanism or obtained as outcome of some physical experiment, one wishes to determine

$$\Pi_{\mathcal{E}_{n+1}}(C) := \text{projection of } C \text{ onto } \mathcal{E}_{n+1}.$$

This problem is handled in [1, 11] by expressing $\mathcal{E}_{n+1}$ as intersection of two "simpler" cones and by applying an alternating projection algorithm. We use here a completely different approach:

- Firstly, we introduce a linear isometry $\phi : \mathbb{S}_\bullet^{n+1} \to \mathbb{S}^n$ and shift the attention to the Loewnerian cone $\mathcal{K} = \phi(\mathcal{E}_{n+1})$. Since ϕ is not merely a linear isomorphism, but also a linear isometry, one has
$$\Pi_{\mathcal{E}_{n+1}}(C) = \phi^T(\Pi_{\mathcal{K}}(A))$$
with $A := \phi(C)$. Indeed,
$$\begin{aligned} \|\phi^T(\Pi_{\mathcal{K}}(A)) - C\| &= \|\Pi_{\mathcal{K}}(A) - A\| = \min_{X\in\mathcal{K}} \|X - A\| \\ &= \min_{W\in\mathcal{E}_{n+1}} \|\phi(W) - A\| = \min_{W\in\mathcal{E}_{n+1}} \|W - C\|. \end{aligned}$$

- Secondly, in order to find the projection of A onto $\mathcal{K}$, we solve the complementarity problem
$$\mathcal{K} \ni X \perp (X - A) \in \mathcal{K}^* \tag{2.41}$$
by using the SNM and a suitable complementarity function.

We now explain the details. As linear isometry $\phi : \mathbb{S}_\bullet^{n+1} \to \mathbb{S}^n$ we consider the map

$$\phi\left(\begin{bmatrix} 0 & b^T \\ b & M \end{bmatrix}\right) = M + \sqrt{2}\,\mathrm{Diag}(b),$$

whose adjoint is given by

$$\phi^T(Z) = \begin{bmatrix} 0 & \frac{1}{\sqrt{2}}[\mathrm{diag}(Z)]^T \\ \frac{1}{\sqrt{2}}\,\mathrm{diag}(Z) & Z^\diamond \end{bmatrix}.$$

Hence, $\mathcal{K} = \phi(\mathcal{E}_{n+1})$ is the Loewnerian cone induced by

$$\mathfrak{F}(U) = \mathfrak{B}(U) - 2U + \sqrt{2}\,\mathrm{Diag}(\mathrm{diag}(U)).$$

A matter of computation shows that

$$\begin{aligned} \mathfrak{F}^{-1}(X) &= (1/\sqrt{8})\,\mathfrak{B}(X) - (1/2)X^\diamond, \\ \mathfrak{F}^T(Y) &= \mathrm{Diag}\left(\sqrt{2}\,\mathrm{diag}(Y) + 2Y^\diamond \mathbf{1}_n\right) - 2Y^\diamond, \\ \mathfrak{F}^{-T}(V) &= (1/\sqrt{2})\,\mathrm{Diag}(V\mathbf{1}_n) - (1/2)V^\diamond. \end{aligned}$$

One can solve the complementarity problem (2.41) by applying the SNM to the system

$$\begin{cases} X - Y - A = \mathbf{0}, \\ \widehat{\kappa}_{\mathrm{fb}}(X, Y) = \mathbf{0}. \end{cases} \tag{2.42}$$

We initialize the algorithm with a random Gaussian matrix X_0 and $Y_0 = X_0 - A$. The complementarity problem (2.41) can also be solved by applying the SNM to

$$\begin{cases} \mathfrak{F}(U) - \mathfrak{F}^{-T}(V) - A = \mathbf{0}, \\ \kappa_{\mathrm{fb}}(U, V) = \mathbf{0}. \end{cases} \tag{2.43}$$

By way of initialization, we generate a random Gaussian matrix Ξ and set $U_0 = \mathfrak{F}^{-1}(\Xi)$ and $V_0 = \mathfrak{F}^{T}(\Xi - A)$. Tables 2.7 and 2.8 report the performance of the SNM applied to (2.42) and (2.43), respectively. The figures displayed in these tables are average values obtained by working with a sample of 10^2 random Gaussian matrices C.

	$n = 25$	$n = 50$	$n = 75$	$n = 100$
success	100%	100%	100%	100%
divergence	0%	0%	0%	0%
ill-conditioning	0%	0%	0%	0%
CPU	5.9	79.6	217.6	839.1

Table 2.7: SNM applied to (2.42).

	$n = 25$	$n = 50$	$n = 75$	$n = 100$
success	100%	100%	100%	100%
divergence	0%	0%	0%	0%
ill-conditioning	0%	0%	0%	0%
CPU	6.2	81.1	222.5	840.6

Table 2.8: SNM applied to (2.43).

Remark 2.26. In a recent paper by Qi [23], the nearest EDM problem is treated by passing to a certain dual optimization problem

$$\text{minimize}_{z \in \mathbb{R}^{n+1}} \ \theta(z) := (1/2) \left\| \Pi_{\mathcal{Q}_{n+1}}[C + \mathrm{Diag}(z)] \right\|^2, \tag{2.44}$$

where

$$\mathcal{Q}_{n+1} := \{A \in \mathbb{S}^{n+1} : \langle x, Ax \rangle \geq 0 \ \text{ when } \langle \mathbf{1}_{n+1}, x \rangle = 0\}.$$

Note that the optimization problem (2.44) is unconstrained. The approach followed in [23] consists in using the SNM to find a root of the gradient map

$$\nabla\theta(z) := \mathrm{diag}\left(\Pi_{\mathcal{Q}_{n+1}}[C + \mathrm{Diag}(z)]\right).$$

Once a root $\bar{z}$ has been found, then one evaluates

$$\Pi_{\mathcal{E}_{n+1}}(C) = \Pi_{\mathcal{Q}_{n+1}}\left(C + \mathrm{Diag}(\bar{z})\right).$$

Projecting onto $\mathcal{Q}_{n+1}$ offers no difficulty.

2.5 By way of conclusion

This work shows that the Semi-Definite Complementarity Problem (SDCP) can be solved efficiently by applying the Semismooth Newton Method (SNM) to the system (2.6). The best results in terms of CPU time are obtained with the help of the Fisher-Burmeister complementarity function κ_{fb}. Some comments concerning the genesis of this work and the related literature are appropriate.

1. The idea of applying the SNM to a cone-constrained eigenvalue problem was considered for the first time by Adly and Seeger [2]. The specific problem treated in [2] is that of finding Pareto-eigenvalues in a given matrix A. This amounts to solve a complementarity problem

$$\begin{cases} Ax - \lambda x - y = \mathbf{0}, \\ \langle \mathbf{1}_n, x \rangle - 1 = 0, \\ \mathbf{0} \leq x \perp y \geq \mathbf{0} \end{cases} \tag{2.45}$$

involving the usual cone $\mathbb{R}^n_+$. Although the model (2.3) looks similar to (2.45), there are important differences. To start with, (2.3) leads to a square system with n^2+n+1 equations, whereas (2.45) leads to a square system with only $2n+1$ equations. Hence, the involved dimensions are not of the same order. Secondly, the nonnegative orthant is a polyhedral cone, whereas the SDP cone is non-polyhedral. As a consequence of this fact, a matrix A has always a finite number of Pareto-eigenvalues, whereas a linear endomorphism $\mathfrak{L}$ on $\mathbb{S}^n$ may perfectly well have a continuum of Loewner-eigenvalues (cf. Proposition 2.27). In other words, the solution sets to (2.3) and (2.45) are structurally different. Thirdly, the squaring technique

$$\begin{aligned} x &= u^{[2]} := u \odot u, \\ y &= v^{[2]} := v \odot v \end{aligned}$$

based on the Hadamard product leads to a companion system

$$\begin{cases} Au^{[2]} - \lambda u^{[2]} - v^{[2]} = \mathbf{0}, \\ \|u\|^2 - 1 = 0, \\ u \odot v = \mathbf{0}, \end{cases} \tag{2.46}$$

that is free of fake solutions. By contrast, the companion system (2.25) associated to (2.3) may well admit fake solutions. In short, a complementarity problem relative to the SDP cone differs substantially from a complementarity problems relative to the nonnegative orthant.

2. The numerical resolution of Example 2.1 is treated here for the first time. Example 2.1 has been the driving motivation behind our work. By contrast, Example 2.2 is mentioned just to illustrate that the SDCP model (2.1) covers a wide variety of applications. The literature dealing with the optimization problem (2.4) is quite extense and comparing different methods for solving (2.4) is beyond the scope of this work.

We also mention that although we have given a list of proper cones that are Loewnerian, a complete characterization of proper cones which are Loewnerian remains as an open question.

Appendix

Let $\sigma(\mathfrak{L}, \mathbb{S}^n_+)$ denote the set of Loewner-eigenvalues of a linear map $\mathfrak{L} : \mathbb{S}^n \to \mathbb{S}^n$. Such set is called the Loewner-spectrum of $\mathfrak{L}$. The following proposition displays a map $\mathfrak{L}$ whose Loewner-spectrum is a set of infinite cardinality.

Proposition 2.27. *Let $\mathfrak{L} : \mathbb{S}^n \to \mathbb{S}^n$ be given by $\mathfrak{L}(X) = \langle C, X\rangle I_n$, where $C \succeq \mathbf{0}$. Then*

$$[\lambda_{\min}(C), \lambda_{\max}(C)] \;\subseteq\; \sigma(\mathfrak{L}, \mathbb{S}^n_+) \;=\; \bigcup_{r=1}^{n} [f_r(C), g_r(C)]\;, \tag{2.47}$$

where $f_r(C)$ and $g_r(C)$ indicate respectively the sum of the r smallest and the sum of the r largest eigenvalues of C.

Proof. Assume that $C \neq \mathbf{0}$, otherwise (2.47) holds trivially. A scalar λ belongs to $\sigma(\mathfrak{L}, \mathbb{S}^n_+)$ if and only if there exists $X \in \mathbb{S}^n$ such that

$$\begin{cases} X \succeq \mathbf{0},\ \mathrm{tr}(X) = 1, \\ \langle C, X\rangle I_n \succeq \lambda X, \\ \langle C, X\rangle = \lambda \|X\|^2. \end{cases}$$

Hence,

$$\sigma(\mathfrak{L}, \mathbb{S}^n_+) = \left\{ \|X\|^{-2} \langle C, X\rangle : X \in \Omega \right\}, \tag{2.48}$$

where Ω stands for the set of matrices $X \in \mathbb{S}^n$ satisfying

$$X \succeq \mathbf{0},\ \mathrm{tr}(X) = 1, \tag{2.49}$$

$$\langle C, X\rangle (\|X\|^2 - \lambda_{\max}(X)) \geq 0. \tag{2.50}$$

Under (2.49), the inequality (2.50) can be written as an equality. The set on the right-hand side of (2.48) remains unchanged if one uses

$$\Omega_0 := \left\{ X \in \mathbb{S}^n : X \succeq \mathbf{0},\ \mathrm{tr}(X) = 1,\ \lambda_{\max}(X) = \|X\|^2 \right\}$$

instead of Ω. On the other hand, one can check that

$$\Omega_0 = \bigcup_{r=1}^{n} \left\{ r^{-1} Q Q^T : Q \in \mathcal{O}(n, r) \right\},$$

where $Q \in \mathcal{O}(n, r)$ indicates that Q is matrix of size $n \times r$ such that $Q^T Q = I_r$. Hence,

$$\sigma(\mathfrak{L}, \mathbb{S}^n_+) = \bigcup_{r=1}^{n} \left\{ \langle C, QQ^T \rangle : Q \in \mathcal{O}(n, r) \right\}.$$

Note that $\langle C, QQ^T \rangle$ ranges over the interval $[f_r(C), g_r(C)]$ as the variable Q ranges over $\mathcal{O}(n, r)$. □

Acknowledgement. Both authors would like to thank the referees for meticulous reading of the manuscript and for several suggestions that improved the presentation.

Bibliography

[1] S. Al-Homidan and H. Wolkowicz. Approximate and exact completion problems for Euclidean distance matrices using semidefinite programming. *Linear Algebra Appl.* 406 (2005), 109–141.

[2] S. Adly and A. Seeger. A nonsmooth algorithm for cone-constrained eigenvalue problems. *Comput. Optim. Appl.*, 49 (2011), 299–318.

[3] D. Azé and J.-B. Hiriart-Urruty. Optimal Hoffman-type estimates in eigenvalue and semidefinite inequality constraints. *J. Global Optim.*, 24 (2002), 133–147.

[4] G.P. Barker and D. Carlson. Cones of diagonally dominant matrices. *Pacific J. Math.*, 57 (1975), 15–32.

[5] A. Barvinok. *A Course in Convexity*. American Mathematical Society, Providence, RI, 2002.

[6] R. Borsdorf and N.J. Higham. A preconditioned Newton algorithm for the nearest correlation matrix. *IMA J. Numer. Anal.*, 30 (2010), 94–107.

[7] X. Chen and P. Tseng. Non-interior continuation methods for solving semidefinite complementarity problems. *Math. Program.* 95 (2003), Ser. A, 431–474.

[8] F.H. Clarke. *Optimization and Nonsmooth Analysis*. John Wiley and Sons, New York, 1983.

[9] P. Gajardo and A. Seeger. Reconstructing a matrix from a partial sampling of Pareto eigenvalues. *Comput. Optim. Appl.*, 51 (2012), 1119–1135.

[10] P. Gajardo and A. Seeger. Solving inverse cone-constrained eigenvalue problems. *Numerische Math.*, 123 (2013), 309–331.

[11] W. Glunt, T.L. Hayden, S. Hong, and J. Wells. An alternating projection algorithm for computing the nearest Euclidean distance matrix. *SIAM J. Matrix Anal. Appl.*, 11 (1990), 589–600.

[12] J.B Hiriart-Urruty and J. Malick. A fresh variational-analysis look at the positive semidefinite matrices world. *J. Optim. Theory Appl.*, 153 (2012), 551–577.

[13] G. Isac, V.A. Bulavsky, V.V. Kalashnikov. *Complementarity, Equilibrium, Efficiency and Economics*. Kluwer Academic Publishers, Dordrecht, 2002.

[14] N.T.B. Kim and D.T. Luc. Normal cones to a polyhedral convex set and generating efficient faces in linear multiobjective programming. *Acta Math. Vietnam.*, 25 (2000), 101–124.

[15] P. Lancaster and L. Rodman. *Algebraic Riccati Equations*. Oxford University Press, New York, 1995.

[16] A.S. Lewis. Eigenvalue-constrained faces. *Linear Algebra Appl.*, 269 (1998), 159–181.

[17] Q. Li, D. Li, and H. Qi. Newton's method for computing the nearest correlation matrix with a simple upper bound. *J. Optim. Theory Appl.*, 147 (2010), 546–568.

[18] Y.J. Liu, L.W. Zhang, and Y.H. Wang. Some properties of a class of merit functions for symmetric cone complementarity problems. *Asia-Pac. J. Oper. Res.*, 23 (2006), 473–495.

[19] J. Malick. A dual approach to semidefinite least-squares problems. *SIAM J. Matrix Anal. Appl.* 26 (2004), 272–284.

[20] J. Malick and H. Sendov. Clarke generalized Jacobian of the projection onto the cone of positive semidefinite matrices. *Set-Valued Anal.*, 14 (2006), 273–293.

[21] R. Nishimura, S. Hayashi, and M. Fukushima. Semidefinite complementarity reformulation for robust Nash equilibrium problems with Euclidean uncertainty sets. *J. Global Optim.* 53 (2012), 107–120.

[22] F. Pukelsheim. *Optimal Design of Experiments*. John Wiley and Sons, Inc., New York, 1993.

[23] H.D. Qi. A semismooth Newton method for the nearest Euclidean distance matrix problem. *SIAM J. Matrix Anal. Appl.*, 34 (2013), 67–93.

[24] L. Qi and J. Sun. A nonsmooth version of Newton's method. *Math. Program.*, 58 (1993), Ser. A, 353–367.

[25] A. Seeger. Eigenvalue analysis of equilibrium processes defined by linear complementarity conditions. *Linear Algebra Appl.* 292 (1999), 1–14.

[26] A. Seeger and D. Sossa. Complementarity problem with respect to Loewnerian cones. *J. Global Optim.*, 2014. Accepted.

[27] R.J. Stern and H. Wolkowicz. Invariant ellipsoidal cones. *Linear Algebra Appl.*, 150 (1991), 81–106.

[28] D. Sun and J. Sun. Semismooth matrix-valued functions. *Math. Oper. Res.* 27 (2002), 150–169.

[29] D. Sun and J. Sun. Strong semismoothness of the Fischer-Burmeister SDC and SOC complementarity functions. *Math. Program.* 103 (2005), Ser. A, 575–581.

[30] L. Zhang, N. Zhang, and L. Pang. Differential properties of the symmetric matrix-valued Fischer-Burmeister function. *J. Optim. Theory Appl.*, 153 (2012), 436–460.

[31] Y. Zhou and M.S. Gowda. On the finiteness of the cone spectrum of certain linear transformations on Euclidean Jordan algebras. *Linear Algebra Appl.* 431 (2009), 772–782.

Chapter 3

Critical angles between two convex cones I. General theory

Alberto Seeger[1] and David Sossa[2]

Abstract. The concept of critical (or principal) angle between two linear subspaces has applications in statistics, numerical linear algebra, and other areas. Such concept has been abundantly studied in the literature, both from a theoretical and computational point of view. Part I of this work is an attempt to build a general theory of critical angles for a pair of closed convex cones. Part II focusses on the practical computation of the maximal angle between specially structured cones.

Mathematics Subject Classification: 15A18, 15A48, 52A40, 90C26, 90C33.
Key words: Maximal angle, critical angle, principal angle, convex cone.

3.1 Introduction

This work is the merging point of two independent sources: the recent theory of critical angles for a closed convex cone, as developed in [6], and the old theory of principal angles

[1]Université d'Avignon, Département de Mathématiques, 33 rue Louis Pasteur, 84000 Avignon, France (alberto.seeger@univ-avignon.fr)

[2]Departamento de Ingeniería Matemática, Centro de Modelamiento Matemático (CNRS UMI 2807), FCFM, Universidad de Chile, Blanco Encalada 2120, Santiago, Chile (dsossa@dim.uchile.cl). This author is supported by CONICYT, Chile.

for a pair of linear subspaces. Let $(\mathbb{X}, \langle\cdot,\cdot\rangle)$ be a Euclidian space[3] of dimension at least two and let $\mathcal{C}(\mathbb{X})$ be the set of nontrivial closed convex cones in $\mathbb{X}$. That a closed convex cone is nontrivial means that it is different from the zero cone and different from the whole space. Computing the maximal angle of a closed convex cone is an issue of importance in a number of applications, see for instance [15]. By definition, the maximal angle of $K \in \mathcal{C}(\mathbb{X})$ is the number

$$\theta_{\max}(K) := \max_{u,v \in K \cap S_{\mathbb{X}}} \arccos\langle u, v\rangle, \tag{3.1}$$

where $S_{\mathbb{X}}$ is the unit sphere of $\mathbb{X}$. By writing down the necessary optimality conditions for the nonconvex optimization problem (3.1), one gets

$$u, v \in K \cap S_{\mathbb{X}}, \quad v - \langle u, v\rangle u \in K^*, \quad u - \langle u, v\rangle v \in K^*, \tag{3.2}$$

where K^* denotes the positive dual cone of K. If the system (3.2) holds, then (u, v) is called a critical pair of K and $\arccos\langle u, v\rangle$ is called a critical angle of K. The study of critical angles in a convex cone was initiated in [6] and further continued in [5, 7, 8, 9, 11]. The purpose of the present work is to build a theory of critical angles for a pair of convex cones. The starting point of our analysis is the formulation of the optimization problem that defines the maximal angle between two convex cones.

Definition 3.1. *Let $P, Q \in \mathcal{C}(\mathbb{X})$. The* maximal angle *of (P, Q) or, more precisely, the maximal angle between P and Q, is given by*

$$\Theta(P, Q) := \max_{u \in P \cap S_{\mathbb{X}},\, v \in Q \cap S_{\mathbb{X}}} \arccos\langle u, v\rangle. \tag{3.3}$$

An antipodal pair *of (P, Q) is any pair $(u, v) \in \mathbb{X}^2$ solving the above maximization problem.*

Antipodal pairs always exist, but they are not unique in general. The formulation of the maximization problem (3.3) is motivated by theoretical and practical considerations.

Example 3.2. Consider the space $\mathbb{S}^n$ of symmetric matrices of order n equipped with the trace inner product $\langle A, B\rangle = \operatorname{tr}(AB)$. An interesting question of linear algebra is to compute the maximal angle between the cones

$$\begin{aligned} \mathbb{S}^n_+ &:= \{A \in \mathbb{S}^n : A \text{ is positive semidefinite}\}, \\ \mathcal{N}_n &:= \{B \in \mathbb{S}^n : B \text{ is nonnegative entrywise}\}. \end{aligned}$$

One usually refers to $\mathbb{S}^n_+$ as the Loewner cone in $\mathbb{S}^n$. By relying on graph theory arguments, Goldberg and Shaker-Monderer [4] obtained a lower bound for the maximal angle $\Theta(\mathbb{S}^n_+, \mathcal{N}_n)$ and proved the asymptotic formula

$$\lim_{n\to\infty} \Theta(\mathbb{S}^n_+, \mathcal{N}_n) = \pi.$$

It remains an open question to compute the exact value of $\Theta(\mathbb{S}^n_+, \mathcal{N}_n)$.

In Section 3.4 we present additional motivations for the study of maximal angles of pairs of convex cones. Next, we write down the necessary optimality conditions for the nonconvex optimization problem (3.3).

[3] It means a finite dimensional inner product space.

Proposition 3.3. *Let $P, Q \in \mathcal{C}(\mathbb{X})$. A necessary condition for $(u, v) \in \mathbb{X}^2$ to be an antipodal pair of (P, Q) is that*

$$\begin{cases} u \in P \cap S_{\mathbb{X}}, \\ v \in Q \cap S_{\mathbb{X}}, \\ v - \langle u, v\rangle u \in P^*, \\ u - \langle u, v\rangle v \in Q^*. \end{cases} \tag{3.4}$$

Proof. Let (u, v) be an antipodal pair of (P, Q). In particular, the component u minimizes the linear form $\langle \cdot, v\rangle$ on $P \cap S_{\mathbb{X}}$. Consider an arbitrary nonzero vector $d \in P$. Clearly,

$$\mathbf{u}(t) := \|u + td\|^{-1}(u + td) \ \in P \cap S_{\mathbb{X}}$$

for all t in some interval $[0, \varepsilon[$. Furthermore, $t = 0$ is a minimum of the function

$$t \in [0, \varepsilon[\mapsto f(t) := \langle \mathbf{u}(t), v\rangle \, .$$

Hence, the right-derivative

$$\lim_{t\to 0^+} \frac{f(t) - f(0)}{t} \ = \ \langle v, d\rangle - \langle u, v\rangle\langle u, d\rangle$$

is nonnegative. This proves the third condition in (3.4). Analogously, the last condition in (3.4) is obtained by using the fact v minimizes $\langle u, \cdot\rangle$ on $Q \cap S_{\mathbb{X}}$. □

In a similar way one can handle the angle minimization problem

$$\begin{cases} \text{mimimize} \ \ \arccos\langle u, v\rangle \\ u \in P \cap S_{\mathbb{X}}, \\ v \in Q \cap S_{\mathbb{X}}, \end{cases}$$

which arises in a number of applications (cf. [14, 19]). The necessary optimality conditions for angle minimization are similar to (3.4), but one must change dual cones by polar cones. In order to avoid repetitions in the exposition, this work focusses only on angle maximization.

Definition 3.4. *Let $P, Q \in \mathcal{C}(\mathbb{X})$.*

(i) A critical pair *of (P, Q) is a pair $(u, v) \in \mathbb{X}^2$ satisfying the system (3.4). The corresponding angle* $\arccos\langle u, v\rangle$ *is called a* critical angle *of (P, Q).*

(ii) A critical pair (u, v) of (P, Q) is called proper *if u and v are not collinear. The corresponding angle is called a* proper critical angle *of (P, Q).*

The main concern of this work is to provide rules for computing antipodal pairs and, more generally, critical pairs. We wish also to analyze the structure of

$$\Gamma(P, Q) := \{\arccos\langle u, v\rangle : (u, v) \text{ satisfies (3.4)}\},$$

a set called the *angular spectrum* of (P, Q). By convention, we write $\Gamma(P, Q) := \{0, \pi\}$ if either P or Q is the whole space $\mathbb{X}$. In general, $\Gamma(P, Q)$ is a nonempty closed subset of the interval $[0, \pi]$. Beware, however, that the cardinality of $\Gamma(P, Q)$ is not necessarily finite. In other words, a pair (P, Q) may have infinitely many critical angles.

3.2 Duality and boundary principles for critical pairs

Let $\Omega_{\mathbb{X}}$ denote the set of all pairs of unit vectors in $\mathbb{X}$ that are not collinear, i.e.,

$$\Omega_{\mathbb{X}} := \{(u,v) \in S_{\mathbb{X}}^2 : |\langle u,v\rangle| \neq 1\}.$$

To each $(u,v) \in \Omega_{\mathbb{X}}$, one can associate its *conjugate pair*

$$g(u,v) = \left(\frac{v - \langle u,v\rangle u}{\sqrt{1-\langle u,v\rangle^2}}\,,\ \frac{u - \langle u,v\rangle v}{\sqrt{1-\langle u,v\rangle^2}} \right).$$

It is not difficult to check that $g : \Omega_{\mathbb{X}} \to \Omega_{\mathbb{X}}$ is a bijection with $g^{-1} = g$. In other words, g is an involution on $\Omega_{\mathbb{X}}$. The following duality principle is an extension of [11, Theorem 2].

Theorem 3.5. *Let $P, Q \in \mathcal{C}(\mathbb{X})$. Let $(u,v) \in \Omega_{\mathbb{X}}$ and $(y,z) \in \Omega_{\mathbb{X}}$ be conjugate pairs. Then (u,v) is a critical pair of (P,Q) if and only if (y,z) is a critical of (P^*,Q^*).*

Proof. Theorem 2 in [11] takes care of the particular case in which P is equal to Q. The proof of the general case follows the same pattern. Assume that (u,v) is critical for (P,Q) and write $\lambda := \langle u,v\rangle$. Clearly,

$$\begin{aligned} y &= [1-\lambda^2]^{-1/2}(v-\lambda u) \in P^* \cap S_{\mathbb{X}}, \\ z &= [1-\lambda^2]^{-1/2}(u-\lambda v) \in Q^* \cap S_{\mathbb{X}}. \end{aligned}$$

Furthermore, $\mu := \langle y,z\rangle = -\lambda$ and

$$\begin{aligned} z - \mu y &= [1-\mu^2]^{1/2}\, u \in P, \\ y - \mu z &= [1-\mu^2]^{1/2}\, v \in Q. \end{aligned}$$

Hence, (y,z) is critical for (P^*,Q^*). The reverse implication is proven in a similar way. □

Corollary 3.6. *Let $P, Q \in \mathcal{C}(\mathbb{X})$. Let $\theta, \psi \in]0,\pi[$ be conjugate angles, i.e., $\theta + \psi = \pi$. Then θ is a critical angle of (P,Q) if and only if ψ is a critical angle of (P^*,Q^*).*

Proof. Suppose that θ is a critical angle of (P,Q). Let (u,v) be any proper critical pair of (P,Q) such that $\cos\theta = \langle u,v\rangle$. Thanks to the duality principle established in Theorem 3.5, the conjugate pair $(y,z) = g(u,v)$ is critical for (P^*,Q^*). Since

$$\langle y,z\rangle = -\langle u,v\rangle = -\cos\theta = \cos(\pi-\theta) = \cos\psi,$$

one deduces that ψ is a critical angle of (P^*,Q^*). The proof of the reverse implication is similar. □

Intuitively speaking, the components u and v of a proper critical pair of (P,Q) should be on the boundaries of P and Q, respectively. The next theorem clarifies this point. For a set C contained in a linear subspace L of $\mathbb{X}$, the symbol $\mathrm{bd}_L(C)$ refers to the boundary of C relative to L.

Theorem 3.7. *Let $L \subseteq \mathbb{X}$ be the smallest linear subspace containing both $P \in \mathcal{C}(\mathbb{X})$ and $Q \in \mathcal{C}(\mathbb{X})$. Suppose that (u, v) is a proper critical pair of (P, Q). Then $u \in \mathrm{bd}_L(P)$ and $v \in \mathrm{bd}_L(Q)$.*

Proof. Since (u, v) is proper, one has $\lambda := \langle u, v\rangle \notin \{-1, 1\}$. Suppose, to the contrary, that u belongs to the interior of P relative to L, i.e., there exists a positive ε such that

$$u + \varepsilon(B_{\mathbb{X}} \cap L) \subseteq P,$$

where $B_{\mathbb{X}}$ is the closed unit ball of $\mathbb{X}$. It follows that

$$0 \;\leq\; \langle v - \lambda u, u + \varepsilon w\rangle \;=\; \varepsilon\langle v - \lambda u, w\rangle$$

for all $w \in B_{\mathbb{X}} \cap L$. The particular choice $w = \|\lambda u - v\|^{-1}(\lambda u - v)$ leads to

$$0 \;\leq\; \varepsilon\|\lambda u - v\|^{-1}\,\langle v - \lambda u, \lambda u - v\rangle \;=\; -\varepsilon\|v - \lambda u\| \;<\; 0,$$

a clear contradiction. This shows that $u \in \mathrm{bd}_L(P)$. The proof of $v \in \mathrm{bd}_L(Q)$ is similar. □

The next corollary follows straightforwardly by combining the duality principle stated in Theorem 3.5 and the boundary principle stated in Theorem 3.7.

Corollary 3.8. *Let $P, Q \in \mathcal{C}(\mathbb{X})$ and (u, v) be a proper critical pair of (P, Q). Then*

$$\begin{cases} v - \langle u, v\rangle u \in \mathrm{bd}_M(P^*), \\ u - \langle u, v\rangle v \in \mathrm{bd}_M(Q^*), \end{cases}$$

where $M \subseteq \mathbb{X}$ is the smallest linear subspace containing both P^* and Q^*.

3.3 Further characterization of criticality and antipodality

Let $\Pi_C(x)$ denote the projection of a point $x \in \mathbb{X}$ onto a nonempty closed convex set $C \subseteq \mathbb{X}$. The next proposition expresses criticality for a pair (P, Q) in terms of the projection maps Π_P and Π_Q.

Proposition 3.9. *Let $P, Q \in \mathcal{C}(\mathbb{X})$. Let u, v be distinct points on the sphere $S_{\mathbb{X}}$. Then (u, v) is a critical pair of (P, Q) if and only if*

$$\begin{cases} \Pi_P(u - v) = (1 - \langle u, v\rangle)\, u, \\ \Pi_Q(v - u) = (1 - \langle u, v\rangle)\, v. \end{cases} \tag{3.5}$$

In particular, a necessary condition for (u, v) to be a critical pair of (P, Q) is that

$$\mathrm{dist}(u - v, P) = \mathrm{dist}(v - u, Q).$$

Proof. Let $\lambda := \langle u, v\rangle$. Let $N_P(u)$ denote the normal cone to P at u. Note that

$$\begin{aligned}\Pi_P(u-v) = (1-\lambda)u &\Leftrightarrow \Pi_P((1-\lambda)^{-1}(u-v)) = u\\ &\Leftrightarrow (1-\lambda)^{-1}(u-v) \in u + N_P(u)\\ &\Leftrightarrow -(v-\lambda u) \in N_P(u)\\ &\Leftrightarrow u \in P,\ v-\lambda u \in P^*.\end{aligned}$$

Similarly, the second condition in (3.5) amounts to saying that $v \in Q$ and $u - \lambda v \in Q^*$. □

Sometimes it is helpful to write the angle maximization problem (3.3) in any of the following equivalent forms

$$\cos[\Theta(P,Q)] = \min_{u\in P\cap S_{\mathbb{X}},\, v\in Q\cap S_{\mathbb{X}}} \langle u, v\rangle, \tag{3.6}$$

$$\kappa(P,Q) := \min_{u\in P\cap S_{\mathbb{X}},\, v\in Q\cap S_{\mathbb{X}}} \left\|\frac{u+v}{2}\right\|. \tag{3.7}$$

The problems (3.3), (3.6), and (3.7) have clearly the same solution set. Furthermore,

$$\kappa(P,Q) = \cos\left[\frac{\Theta(P,Q)}{2}\right]. \tag{3.8}$$

The next theorem relates (3.3) to the minimization problem

$$\chi(P,Q) := \min_{z\in S_{\mathbb{X}}} \max\{\operatorname{dist}(z,P), \operatorname{dist}(z,-Q)\}. \tag{3.9}$$

Although it is not clear at first sight, it turns out that solving (3.3) is equivalent to solving (3.9).

Theorem 3.10. *Let $P, Q \in \mathcal{C}(\mathbb{X})$. Then*

$$\Theta(P,Q) = 2\arccos[\chi(P,Q)]. \tag{3.10}$$

Suppose, in addition, that P, Q are not equal to a common ray. In such a case, the solution set $\mathcal{A}(P,Q)$ to the angle maximization problem (3.3) and the solution set $\mathcal{R}(P,Q)$ to the problem (3.9) are related as follows:

$$\mathcal{R}(P,Q) = \left\{\frac{u-v}{\|u-v\|} : (u,v) \in \mathcal{A}(P,Q)\right\}, \tag{3.11}$$

$$\mathcal{A}(P,Q) = \left\{\left(\frac{\Pi_P(z)}{\|\Pi_P(z)\|}, \frac{\Pi_Q(-z)}{\|\Pi_Q(-z)\|}\right) : z \in \mathcal{R}(P,Q)\right\}. \tag{3.12}$$

Proof. If P and Q are equal to a common ray, then both sides of (3.10) are equal to 0. Suppose now that P and Q are not equal to a common ray. Suppose also that $P \cap -Q = \{0\}$, otherwise both sides of (3.10) are equal to π and the proof of (3.11)-(3.12) is immediate. Let z_0 be a solution to (3.9). Hence,

$$\begin{aligned}\operatorname{dist}(z_0,P) &= \operatorname{dist}(z_0,-Q) = \chi(P,Q),\\ \|\Pi_P(z_0)\| &= \|\Pi_Q(-z_0)\| = s,\end{aligned}$$

with $s := (1 - [\chi(P,Q)]^2)^{1/2}$ belonging to $]0,1[$. The pair

$$\begin{aligned}(u_0, v_0) &:= \big(\|\Pi_P(z_0)\|^{-1}\Pi_P(z_0)\,,\, \|\Pi_Q(-z_0)\|^{-1}\Pi_Q(-z_0) \big) \\ &= (1/s)\,(\,\Pi_P(z_0)\,,\,\Pi_Q(-z_0)\,)\end{aligned}$$

is then well defined. We claim that

$$z_0 = \|u_0 - v_0\|^{-1}(u_0 - v_0), \tag{3.13}$$

$$\kappa(P,Q) \le \left\| \frac{u_0 + v_0}{2} \right\| = \chi(P,Q). \tag{3.14}$$

The inequality in (3.14) is obvious, but it is added for convenience. Let $c : \mathbb{X} \to \mathbb{R}$ be the objective function of the minimization problem (3.9), i.e.,

$$c(z) := \max\{\mathrm{dist}(z,P), \mathrm{dist}(z,-Q)\}.$$

Since z_0 minimizes $(1/2)c^2(\cdot)$ on $S_{\mathbb{X}}$, it satisfies the optimality condition

$$\lambda_1\,(z_0 - \Pi_P(z_0)) + \lambda_2\,(z_0 - \Pi_{-Q}(z_0)) + \mu z_0 = 0, \tag{3.15}$$

where $\mu \in \mathbb{R}$ is a Lagrange multiplier and $\lambda_1, \lambda_2 \in \mathbb{R}$ are nonnegative scalars adding up to 1. Since $z_0 - \Pi_P(z_0)$ is orthogonal to $\Pi_P(z_0)$ and $z_0 - \Pi_{-Q}(z_0)$ is orthogonal to $\Pi_{-Q}(z_0)$, one gets

$$\langle z_0, \Pi_P(z_0)\rangle = \langle z_0, \Pi_{-Q}(z_0)\rangle = s^2.$$

This and (3.15) yield $\mu + 1 = s^2$ and $\lambda_1 = \lambda_2 = 1/2$. Hence,

$$z_0 = (1/2s^2)\,(\Pi_P(z_0) + \Pi_{-Q}(z_0)) = (1/2s)(u_0 - v_0).$$

This proves (3.13) and shows that

$$z_0 \in \mathrm{cone}\,\{\Pi_P(z_0), \Pi_{-Q}(z_0)\} \subseteq L := \mathrm{span}\{u_0, v_0\}.$$

In the plane L, consider the rectangular triangles $\mathrm{co}\{0, z_0, \Pi_P(z_0)\}$ and $\mathrm{co}\{0, z_0, \Pi_{-Q}(z_0)\}$. Both triangles have the same angle at the vertex 0, namely, $\psi = \arcsin[\chi(P,Q)]$. It is plain to see that

$$\langle u_0, -v_0\rangle = \cos(2\psi) = 1 - 2\sin^2\psi = 1 - 2[\chi(P,Q)]^2.$$

This leads directly to the equality (3.14). Next, let (u_1, v_1) be an antipodal pair of (P,Q). Then $z_1 := \|u_1 - v_1\|^{-1}(u_1 - v_1)$ is well defined. From Proposition 3.9 one knows that

$$\begin{cases} \Pi_P(z_1) = \|u_1 - v_1\|^{-1}(1 - \langle u_1, v_1\rangle)\,u_1, \\ \Pi_Q(-z_1) = \|u_1 - v_1\|(1 - \langle u_1, v_1\rangle)\,v_1. \end{cases} \tag{3.16}$$

Hence,

$$\mathrm{dist}(z_1, P) = \|z_1 - \Pi_P(z_1)\| = \frac{\|v_1 - \langle u_1, v_1\rangle u_1\|}{\|u_1 - v_1\|} = \left\| \frac{u_1 + v_1}{2} \right\| = \kappa(P,Q)$$

and, similarly, $\mathrm{dist}(z_1, -Q) = \mathrm{dist}(-z_1, Q) = \kappa(P,Q)$. It follows that

$$\chi(P,Q) \le c(z_1) = \kappa(P,Q). \tag{3.17}$$

From (3.16) one deduces also that

$$(u_1, v_1) \;=\; \big(\|\Pi_P(z_1)\|^{-1}\Pi_P(z_1), \|\Pi_Q(-z_1)\|^{-1}\Pi_Q(-z_1)\big)\,. \tag{3.18}$$

The combination of (3.13)-(3.14) and (3.17)-(3.18) completes the proof of the theorem.

□

3.4 Antipodality, pointedness and reproducibility

The sum of two closed convex cones may not be closed. The next proposition is part of the folklore on convex cones, cf. [2, Theorem 3.2]. A pair (P, Q) of elements in $\mathcal{C}(\mathbb{X})$ is said to be *pointed* if $P \cap -Q = \{0\}$. A single cone $K \in \mathcal{C}(\mathbb{X})$ is declared pointed if the pair (K, K) is pointed.

Proposition 3.11. *Let $P, Q \in \mathcal{C}(\mathbb{X})$. The following conditions are equivalent and imply that $P + Q$ is closed:*

(a) (P, Q) is pointed.

(b) There exists a positive constant β such that

$$\beta(\|u\| + \|v\|) \leq \|u + v\| \quad \text{for all } u \in P, v \in Q. \tag{3.19}$$

The reverse triangular inequality (3.19) holds of course with $\beta = 0$, but such choice is useless. What is interesting to know is the best constant

$$\beta(P, Q) := \max\{\beta \in [0, 1] : \beta \text{ satisfies (3.19)}\}.$$

Such a coefficient measures to which extent the pair (P, Q) is pointed. The next proposition relates $\beta(P, Q)$ to the maximal angle of (P, Q).

Proposition 3.12. *Let $P, Q \in \mathcal{C}(\mathbb{X})$. Then*

$$\beta(P, Q) = \cos\left[\frac{\Theta(P, Q)}{2}\right].$$

Proof. We need to prove that $\beta(P, Q) = \kappa(P, Q)$. Assume that P and Q are not equal to a common ray, otherwise we are done. Clearly,

$$\beta(P, Q) = \min_{\substack{u \in P, v \in Q \\ (u,v) \neq (0,0)}} \frac{\|u + v\|}{\|u\| + \|v\|}.$$

But the Dunkl-Williams inequality implies that

$$\frac{1}{2}\left\|\frac{u}{\|u\|} + \frac{u}{\|u\|}\right\| \leq \frac{\|u + v\|}{\|u\| + \|v\|}$$

whenever $u, v \in \mathbb{X}$ are nonzero vectors. Hence, $\beta(P, Q)$ is greater than or equal to $\kappa(P, Q)$. The reverse inequality is obvious. □

A pair (P, Q) of elements in $\mathcal{C}(\mathbb{X})$ is said to be *reproducing* if $P - Q = \mathbb{X}$. A single cone $K \in \mathcal{C}(\mathbb{X})$ is called reproducing if the pair (K, K) is reproducing. Clearly, (P, Q) is reproducing if and only if (P^*, Q^*) is pointed. The next result comes then without surprise.

Proposition 3.13. *Let $P, Q \in \mathcal{C}(\mathbb{X})$. The following conditions are equivalent and imply that $P^* + Q^*$ is closed:*

(a) (P, Q) is reproducing.

(b) There exists a positive constant α such that

$$\alpha B_{\mathbb{X}} \subseteq \mathrm{co}((P \cup -Q) \cap B_{\mathbb{X}}). \tag{3.20}$$

The notation "co" refers of course to the convex hull operation. One can see Proposition 3.13 as a sort of dual version of Proposition 3.11. The coefficient

$$\alpha(P, Q) := \max\{\alpha \in [0, 1] : \alpha \text{ satisfies (3.20)}\}$$

measures to which extent the pair (P, Q) is reproducing. The next proposition shows that evaluating the reproducibility coefficient of (P, Q) amounts to compute the maximal angle of (P^*, Q^*).

Proposition 3.14. *Let $P, Q \in \mathcal{C}(\mathbb{X})$. Then*

$$\alpha(P, Q) = \cos\left[\frac{\Theta(P^*, Q^*)}{2}\right].$$

Proof. By using duality arguments (namely, calculus rules for polar sets) one can show that the inclusion (3.20) can be written in the equivalent form

$$\alpha\left[(B_{\mathbb{X}} + P^*) \cap (B_{\mathbb{X}} - Q^*)\right] \subseteq B_{\mathbb{X}}.$$

Hence, $\alpha(P^*, Q^*)$ is equal to the coefficient

$$\nu(P, Q) := \max\{r \in [0, 1] : r\left[(B_{\mathbb{X}} + P) \cap (B_{\mathbb{X}} - Q)\right] \subseteq B_{\mathbb{X}}\}.$$

By proceeding as in [10, Theorem 2], one can check that $\nu(P, Q) = \chi(P, Q)$. Theorem 3.10 does the rest of the job. □

3.5 Lipschitzness of the maximal angle function

Topological issues on $\mathcal{C}(\mathbb{X})$ are relative to the spherical metric δ, which is defined by

$$\delta(K_1, K_2) := \mathrm{haus}(K_1 \cap S_{\mathbb{X}}, K_2 \cap S_{\mathbb{X}}).$$

Here,

$$\mathrm{haus}(C_1, C_2) := \max\left\{\max_{x \in C_1} \mathrm{dist}(x, C_2), \max_{x \in C_2} \mathrm{dist}(x, C_1)\right\}$$

stands for the classical Pompeiu-Hausdorff distance between a pair C_1, C_2 of nonempty compact subsets of $\mathbb{X}$. Convergence with respect to the spherical metric is equivalent to

convergence in the Painlevé-Kuratowski sense. Topological issues on the product space $\mathcal{C}^2(\mathbb{X}) := \mathcal{C}(\mathbb{X}) \times \mathcal{C}(\mathbb{X})$ refer to the metric

$$\Delta((P_1, Q_1), (P_2, Q_2)) := \max\{\delta(P_1, P_2), \delta(Q_1, Q_2)\}.$$

The next proposition concerns the continuity behavior of the multivalued map Γ. Upper and lower-semicontinuity of multivalued maps between metric spaces are understood in the classical sense, cf. [1, Section 1.4].

Proposition 3.15. *The multivalued map $\Gamma : \mathcal{C}^2(\mathbb{X}) \rightrightarrows \mathbb{R}$ is upper-semicontinuous, but not lower-semicontinuous.*

Proof. The values of Γ are closed subsets of the compact interval $[0, \pi]$. For proving that Γ is upper-semicontinuous, it is enough to check that

$$\operatorname{gr}\Gamma := \{(P, Q, \theta) \in \mathcal{C}^2(\mathbb{X}) \times \mathbb{R} : \theta \in \Gamma(P, Q)\}$$

is a closed set. Let $\{(P_k, Q_k, \theta_k)\}_{k\in\mathbb{N}}$ be a sequence in $\operatorname{gr}\Gamma$ converging to some $(P, Q, \theta) \in \mathcal{C}^2(\mathbb{X}) \times \mathbb{R}$. For each $k \in \mathbb{N}$, there exists a pair $(u_k, v_k) \in \mathbb{X}^2$ such that

$$\begin{cases} \theta_k = \arccos\langle u_k, v_k\rangle, \\ u_k \in P_k \cap S_{\mathbb{X}}, \\ v_k \in Q_k \cap S_{\mathbb{X}}, \\ v_k - \langle u_k, v_k\rangle u_k \in P_k^*, \\ u_k - \langle u_k, v_k\rangle v_k \in Q_k^*. \end{cases} \tag{3.21}$$

Let (u, v) be the limit of some subsequence $\{(u_{\varphi(k)}, v_{\varphi(k)})\}_{k\in\mathbb{N}}$. We write (3.21) with $\varphi(k)$ instead of k. By passing then to the limit, one deduces that $(P, Q, \theta) \in \operatorname{gr}\Gamma$. We now prove that Γ is not lower-semicontinuous. Let $e_1, e_2 \in S_{\mathbb{X}}$ be orthogonal. For each integer $k \geq 1$, let

$$\begin{aligned} u_k &:= -k^{-1} e_1 + \left(1 - k^{-2}\right)^{1/2} e_2, \\ P_k &= \mathbb{R}_+ u_k := \{t u_k : t \geq 0\}, \\ Q_k &= H_{e_1} := \{x \in \mathbb{X} : \langle e_1, x\rangle \geq 0\}. \end{aligned}$$

A matter of computation shows that $(u_k, -u_k)$ is the unique critical pair of (P_k, Q_k). Thus, $\Gamma(P_k, Q_k) = \{\pi\}$. On the other hand,

$$\Gamma\left(\lim_{k\to\infty} P_k, \lim_{k\to\infty} Q_k\right) = \Gamma(\mathbb{R}_+ e_2, H_{e_1}) = \{0, \pi\}.$$

This proves that Γ is not lower-semicontinuous. □

As shown in the next theorem, the maximal angle function $\Theta : \mathcal{C}^2(\mathbb{X}) \to \mathbb{R}$ is not merely continuous, but it is also Lipschitzian.

Theorem 3.16. *There exists a constant $\ell_{\mathbb{X}}$ such that*

$$|\Theta(P_1, Q_1) - \Theta(P_2, Q_2)| \;\leq\; \ell_{\mathbb{X}}\, \Delta((P_1, Q_1), (P_2, Q_2))$$

for all $(P_1, Q_1), (P_2, Q_2) \in \mathcal{C}^2(\mathbb{X})$.

Proof. One knows from (3.8) that $\Theta : \mathcal{C}^2(\mathbb{X}) \to \mathbb{R}$ admits the characterization

$$\Theta(P, Q) = 2 \arccos\left(\kappa(P, Q)\right). \tag{3.22}$$

We claim that κ satisfies the Lipschitz condition

$$|\kappa(P_1, Q_1) - \kappa(P_2, Q_2)| \leq \Delta((P_1, Q_1), (P_2, Q_2)). \tag{3.23}$$

The proof of (3.23) follows the same pattern as in [10, Lemma 1]. Let $u_2 \in P_2 \cap S_{\mathbb{X}}$ and $v_2 \in Q_2 \cap S_{\mathbb{X}}$ be such that $2\kappa(P_2, Q_2) = \|u_2 + v_2\|$. Let u_1, v_1 be the projections of u_2, v_2 onto $P_1 \cap S_{\mathbb{X}}$ and $Q_1 \cap S_{\mathbb{X}}$, respectively. Hence,

$$\begin{aligned} 2(\kappa(P_1, Q_1) - \kappa(P_2, Q_2)) &\leq \|u_1 + v_1\| - \|u_2 + v_2\| \\ &\leq \|u_1 - u_2\| + \|v_1 - v_2\| \\ &= \operatorname{dist}(u_2, P_1 \cap S_{\mathbb{X}}) + \operatorname{dist}(v_2, Q_1 \cap S_{\mathbb{X}}) \\ &\leq \mathrm{e}(P_2, P_1) + \mathrm{e}(Q_2, Q_1), \end{aligned}$$

where one uses the notation

$$\mathrm{e}(K_2, K_1) := \sup_{u \in K_2 \cap S_{\mathbb{X}}} \operatorname{dist}(u, K_1 \cap S_{\mathbb{X}}).$$

In a similar way one gets

$$2(\kappa(P_2, Q_2) - \kappa(P_1, Q_1)) \leq \mathrm{e}(P_1, P_2) + \mathrm{e}(Q_1, Q_2).$$

Thus,

$$\begin{aligned} 2|\kappa(P_1, Q_1) - \kappa(P_2, Q_2)| &\leq \max\{\mathrm{e}(P_2, P_1) + \mathrm{e}(Q_2, Q_1), \mathrm{e}(P_1, P_2) + \mathrm{e}(Q_1, Q_2)\} \\ &\leq \underbrace{\max\{\mathrm{e}(P_2, P_1), \mathrm{e}(P_1, P_2)\}}_{\delta(P_1, P_2)} + \underbrace{\max\{\mathrm{e}(Q_2, Q_1), \mathrm{e}(Q_1, Q_2)\}}_{\delta(Q_1, Q_2)}. \end{aligned}$$

This leads to (3.23). Next we observe that

$$\Theta(P, Q) = \arccos\left(1 - (1/2)[\operatorname{diam}(P, Q)]^2\right), \tag{3.24}$$

where

$$\operatorname{diam}(P, Q) := \max_{u \in P \cap S_{\mathbb{X}},\, v \in Q \cap S_{\mathbb{X}}} \|u - v\|.$$

It is not difficult to check that

$$\begin{aligned} |\operatorname{diam}(P_1, Q_1) - \operatorname{diam}(P_2, Q_2)| &\leq \delta(P_1, P_2) + \delta(Q_1, Q_2) \\ &\leq 2\,\Delta((P_1, Q_1), (P_2, Q_2)). \end{aligned} \tag{3.25}$$

The Lipschitzness of Θ is obtained by combining (3.22), (3.23), (3.24), and (3.25). To see this, one can follow the same procedure as in [17, Theorem 2]. The details are omitted. □

3.6 Critical angles in a pair of linear subspaces

Which is the minimal angle between a pair of linear subspaces? And which one is the maximal angle? Are they other interesting angles, besides the minimal and the maximal one? This sort of questions has lead to develop the classical theory of principal angles. Recall that the principal angles $\theta_1, \ldots, \theta_m$ of a pair (P, Q) of nontrivial linear subspaces of $\mathbb{X}$ are defined recursively by

$$\cos\theta_k = \max_{u \in P_k \cap S_{\mathbb{X}},\, v \in Q_k \cap S_{\mathbb{X}}} \langle u, v\rangle, \tag{3.26}$$

where $m := \min\{\dim P, \dim Q\}$ and

$$\begin{cases} P_1 := P,\ Q_1 := Q, \\ P_{k+1} := \{x \in P_k : \langle u_k, x\rangle = 0\}, \\ Q_{k+1} := \{x \in Q_k : \langle v_k, x\rangle = 0\}, \\ (u_k, v_k) \text{ solution to (3.26)}. \end{cases}$$

The vectors u_k and v_k are not uniquely defined, but the θ_k are unique. Interesting material on principal angles can be found in the linear algebra book by Meyer [12, Section 5.15], see also the references [3, 13, 16]. When P and Q are nontrivial linear subspaces of $\mathbb{X}$, the system (3.4) becomes

$$\begin{cases} u \in P \cap S_{\mathbb{X}}, \\ v \in Q \cap S_{\mathbb{X}}, \\ v - \langle u, v\rangle u \in P^{\perp}, \\ u - \langle u, v\rangle v \in Q^{\perp}, \end{cases} \tag{3.27}$$

where $\perp$ indicates orthogonal complementation relative to $\mathbb{X}$. In this special context, there is no distinction between criticality for angle maximization and criticality for angle minimization. As a first elementary observation, we mention the following conjugacy principle.

Proposition 3.17. *Let P and Q be nontrivial linear subspaces of $\mathbb{X}$. Let $\theta, \psi \in [0, \pi]$ be conjugate angles. Then θ is a critical angle of (P, Q) if and only if ψ is a critical angle of (P, Q).*

Proof. Clearly, (u, v) satisfies (3.27) if and only if $(u, -v)$ satisfies (3.27). It suffices now to observe that $\arccos\langle u, -v\rangle$ and $\arccos\langle u, v\rangle$ are conjugate angles. □

By the way, the combination of Corollary 3.6 and Proposition 3.17 yields a duality result established by Miao and Ben-Israel [13, Theorem 3]. In view of Proposition 3.17, it is enough to compute the critical angles of (P, Q) that are in the subinterval $[0, \pi/2]$. The remaining critical angles are obtained by conjugation. The next theorem shows that the principal angles of (P, Q) are equal to the critical angles of (P, Q) that are in $[0, \pi/2]$. In what follows, we use the notation

$$\mathcal{O}(\mathbb{R}^n, \mathbb{X}) := \{W \in \mathcal{L}(\mathbb{R}^n, \mathbb{X}) : W^T W = I_n\},$$

where I_n is the identity matrix of order n and $\mathcal{L}(\mathbb{R}^n, \mathbb{X})$ is the vector space of linear maps from $\mathbb{R}^n$ to $\mathbb{X}$. The symbol $\mathrm{Im}W$ refers to the image space of W.

Theorem 3.18. *Let $P = \operatorname{Im} U$ and $Q = \operatorname{Im} V$ be nontrivial linear subspaces of $\mathbb{X}$ represented by $U \in \mathcal{O}(\mathbb{R}^p, \mathbb{X})$ and $V \in \mathcal{O}(\mathbb{R}^q, \mathbb{X})$, respectively. For $\theta \in [0, \pi/2]$, the following four conditions are equivalent:*

(a) θ is a critical angle of (P, Q).

(b) θ is a principal angle of (P, Q).

(c) $\cos\theta$ is a singular value of the rectangular matrix $E := V^T U$.

(d) There are unit vectors $x \in \mathbb{R}^p$ and $y \in \mathbb{R}^q$ such that

$$\begin{bmatrix} 0 & E^T \\ E & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \cos\theta \begin{bmatrix} x \\ y \end{bmatrix}. \tag{3.28}$$

Furthermore, if x and y are as in (d), then (Ux, Vy) is a critical pair of (P, Q) and $\cos\theta = \langle Ux, Vy \rangle$.

Proof. (b) $\Leftrightarrow$ (c). This equivalence is stated in [3, Theorem 1].
(a) $\Rightarrow$ (d). Let the angle θ be formed with some pair (u, v) satisfying the system (3.27). There are unit vectors $x \in \mathbb{R}^p$ and $y \in \mathbb{R}^q$ such that $(u, v) = (Ux, Vy)$. Hence, $\cos\theta = \langle u, v \rangle = \langle Ux, Vy \rangle$ and

$$\begin{cases} Vy - (\cos\theta)Ux \in P^\perp, \\ Ux - (\cos\theta)Vy \in Q^\perp. \end{cases}$$

Since $P^\perp = \operatorname{Ker}(U^T)$ and $Q^\perp = \operatorname{Ker}(V^T)$, one gets

$$\begin{cases} E^T y = (\cos\theta)x, \\ Ex \;\; = (\cos\theta)y. \end{cases} \tag{3.29}$$

which is an equivalent way of writing (3.28).
(d) $\Rightarrow$ (c). By exchanging the roles of P and Q if necessary, one may suppose that $\min\{p, q\} = p$. Let x and y be as in (d). From (3.29), one gets $E^T E x = (\cos\theta)^2 x$. Hence, $(\cos\theta)^2$ is an eigenvalue of $E^T E$ and $\cos\theta$ is a singular value of E.
(c) $\Rightarrow$ (a). One can write $E = F \Sigma G^T$, where $G = [g_1 \dots g_p]$ and $F = [f_1 \dots f_q]$ are orthogonal matrices of order p and q, respectively, and Σ is a $q \times p$ diagonal matrix with the singular values of E placed on the diagonal entries. Let $\cos\theta$ be a singular value of E. Suppose that $\cos\theta$ is placed on k-th diagonal entry of Σ. One gets $E g_k = (\cos\theta) f_k$ and $E^T f_k = (\cos\theta) g_k$. Hence, the system (3.28) holds with $x = g_k$ and $y = f_k$. One deduces that $(u, v) = (U g_k, V f_k)$ is a critical pair of (P, Q) producing the angle θ. □

Let P and Q be as in Theorem 3.18. By combining Theorem 3.18 and Proposition 3.17, one obtains

$$\Gamma(P, Q) = \bigcup_{\sigma \in \Sigma(E)} \{\arccos\sigma, \pi - \arccos\sigma\},$$

where $\Sigma(E)$ is the set of singular values of E. In particular, the critical angles of (P, Q) are at most $2\min\{p, q\}$ and they come in conjugate pairs. Such upper bound can be sharpened as follows.

Proposition 3.19. *Let P and Q be nontrivial linear subspaces of $\mathbb{X}$ of dimensions p and q, respectively. Let r be the dimension of $P \cap Q$. Then*

$$\operatorname{card}\left[\Gamma(P,Q)\right] \leq 2\min\{p,q\} - 2\max\{0, r-1\}. \tag{3.30}$$

Proof. We assume that $1 \leq r < \min\{p,q\}$, otherwise we are done. One can represent P and Q as in Theorem 3.18, with the additional feature that $U = [W, U_0]$ and $V = [W, V_0]$ have a portion in common. The common part $W \in \mathcal{O}(\mathbb{R}^r, \mathbb{X})$ serves to represent the intersection of P and Q, i.e., $P \cap Q = \mathrm{Im} W$. One has

$$\begin{aligned} &W^T W = I_r,\ \ U_0^T U_0 = I_{p-r},\ V_0^T V_0 = I_{q-r}, \\ &W^T U_0 = 0,\ U_0^T W = 0, W^T V_0 = 0,\ V_0^T W = 0, \end{aligned}$$

Let the angle θ be formed with some pair (u,v) satisfying the system (3.27). There are vectors $\xi, \eta \in \mathbb{R}^r$ and $x \in \mathbb{R}^{p-r}$, $y \in \mathbb{R}^{q-r}$ such that $\|\xi\|^2 + \|x\|^2 = 1$, $\|\eta\|^2 + \|y\|^2 = 1$, and

$$(u,v) = (W\xi + U_0 x, W\eta + V_0 y). \tag{3.31}$$

By substituting (3.31) into (3.27), one gets

$$\left\{ \begin{array}{rcl} W^T[W\eta + V_0 y - \lambda(W\xi + U_0 x)] &=& 0, \\ U_0^T[W\eta + V_0 y - \lambda(W\xi + U_0 x)] &=& 0, \\ W^T[W\xi + U_0 x - \lambda(W\eta + V_0 y)] &=& 0, \\ V_0^T[W\xi + U_0 x - \lambda(W\eta + V_0 y)] &=& 0, \end{array} \right.$$

with $\lambda = \langle W\xi + U_0 x, W\eta + V_0 y\rangle$. After simplification, one obtains $\eta = \lambda\xi$, $\xi = \lambda\eta$, and

$$\left\{ \begin{array}{rcl} U_0^T V_0 y &=& \lambda x, \\ V_0^T U_0 x &=& \lambda y. \end{array} \right. \tag{3.32}$$

If $\lambda \notin \{-1, 1\}$, then $\xi = 0$, $\eta = 0$ and $x \in \mathbb{R}^{p-q}, y \in \mathbb{R}^{q-r}$ are unit vectors satisfying (3.32). Hence, λ may take at most $2\min\{p-r, q-r\}$ different values. To this count one should add the potential candidates $\lambda = -1$ and $\lambda = 1$. One gets in this way the upper estimate

$$\operatorname{card}\left[\Gamma(P,Q)\right] \leq 2\min\{p-r, q-r\} + 2.$$

This proves (3.30). □

3.7 Critical angles in a pair of polyhedral cones

This section is devoted to the analysis of critical angles in a pair (P,Q) of polyhedral cones. We suppose that the reader is acquainted with the theory of faces of convex polyhedra. The notation that we use is as follows:

$$\begin{array}{rcl} \mathcal{F}(P) &:=& \{F \subseteq \mathbb{X} : F \text{ is a nonzero face of } P\}, \\ \mathrm{span}F &:=& \text{linear subspace spanned by } F, \\ \mathrm{ri}F &:=& \text{relative interior of } F, \\ \dim F &:=& \text{dimension of } \mathrm{span}F, \\ \Pi^F &:=& \text{orthogonal projector onto } \mathrm{span}F. \end{array}$$

For a nonzero vector u in a polyhedral cone $P \in \mathcal{C}(\mathbb{X})$, there exists a unique $F \in \mathcal{F}(P)$ such that $u \in \mathrm{ri}F$. Such F is called the face associated to u.

Theorem 3.20. *Let $P, Q \in \mathcal{C}(\mathbb{X})$ be polyhedral cones. If (u, v) is a critical pair of (P, Q), then*

$$\Pi^F v = \langle u, v\rangle u, \quad \Pi^E u = \langle u, v\rangle v, \tag{3.33}$$

where F is the face of P associated to u and E is the face of Q associated to v. In particular,

$$\Gamma(P, Q) \subseteq \bigcup_{F\in\mathcal{F}(P)} \bigcup_{E\in\mathcal{F}(Q)} \Gamma(\mathrm{span}F, \mathrm{span}E). \tag{3.34}$$

Proof. By assumption, $u \in \mathrm{ri}F$ and $v \in \mathrm{ri}E$ satisfy the criticality conditions stated in (3.4). By proceeding as [18, Theorem 3.4], one can check that

$$\begin{cases} v - \langle u, v\rangle u \in (\mathrm{span}F)^\perp, \\ u - \langle u, v\rangle v \in (\mathrm{span}E)^\perp. \end{cases}$$

But this is clearly equivalent to (3.33). □

By using the inclusion (3.34), one gets the upper bound

$$\mathrm{card}[\Gamma(P, Q)] \leq \sum_{F\in\mathcal{F}(P)} \sum_{E\in\mathcal{F}(Q)} \mathrm{card}[\Gamma(\mathrm{span}F, \mathrm{span}E)]. \tag{3.35}$$

The above inequality becomes an equality, for instance, if P and Q are nontrivial linear subspaces. Since the double sum in (3.35) is finite, any pair of polyhedral cones has finitely many critical angles. By combining (3.35) and the estimate

$$\mathrm{card}[\Gamma(\mathrm{span}F, \mathrm{span}E)] \leq 2\min\{\dim F, \dim E\},$$

one gets in particular

$$\mathrm{card}\,[\Gamma(P, Q)] \leq 2 \sum_{k=1}^{\dim P} \sum_{\ell=1}^{\dim Q} c_P(k) c_Q(\ell) \min\{k, \ell\}, \tag{3.36}$$

where $c_P(k)$ stands for the number of k-dimensional faces of P. The upper bound (3.36) is coarse in general, so we shall not elaborate further on the practical evaluation of such an expression.

Corollary 3.21. *Let $P, Q \in \mathcal{C}(\mathbb{X})$ be polyhedral cones. Let (u_1, v) and (u_2, v) be critical pairs of (P, Q). If u_1 and u_2 have the same associated face, then the critical angles $\theta_1 := \arccos\langle u_1, v\rangle$ and $\theta_2 := \arccos\langle u_2, v\rangle$ are equal or conjugate.*

Proof. Suppose that u_1 and u_2 have $F \in \mathcal{F}(P)$ as common associated face. In such a case, Theorem 3.20 yields

$$\langle u_1, v\rangle u_1 = \Pi^F v = \langle u_2, v\rangle u_2.$$

By taking norms, one sees that $\langle u_1, v\rangle$ and $\langle u_2, v\rangle$ have the same absolute value. This proves that θ_1 and θ_2 are equal or conjugate. □

We now concentrate on the numerical computation of the critical angles of a pair

$$(P, Q) = (G(\mathbb{R}^p_+), H(\mathbb{R}^q_+)) \tag{3.37}$$

of polyhedral cones in $\mathbb{R}^n$. Here, $G = [g_1, \ldots, g_p]$ and $H = [h_1, \ldots, h_q]$ are matrices of size $n \times p$ and $n \times q$, respectively. Without loss of generality, we assume that

$$\begin{cases} g_1, \ldots, g_p \text{ are conically independent unit vectors,} \\ h_1, \ldots, h_q \text{ are conically independent unit vectors.} \end{cases} \tag{3.38}$$

That a collection of vectors is conically independent simply means that no element from the collection can be expressed as positive linear combination of those remaining. For notational convenience, we introduce the index sets

$$\begin{aligned} \mathcal{I}(G) &:= \{I \subseteq \{1, \ldots, p\} : I \neq \emptyset \text{ and } \{g_i : i \in I\} \text{ is linearly independent}\}, \\ \mathcal{J}(H) &:= \{J \subseteq \{1, \ldots, q\} : J \neq \emptyset \text{ and } \{h_j : j \in J\} \text{ is linearly independent}\}. \end{aligned}$$

The cardinality of an index set, say I, is denoted by $|I|$. We write G_I to indicate the submatrix of G with columns indexed by I. The definition of H_J is similar. Without further ado, we state the next theorem.

Theorem 3.22. *Let (P, Q) be as in (3.37)-(3.38). Then the following statements are equivalent:*

(a) $\theta \in \Gamma(P, Q)$,

(b) there are index sets $I \in \mathcal{I}(G)$, $J \in \mathcal{J}(H)$ and vectors $\xi \in \mathbb{R}^{|I|}$, $\eta \in \mathbb{R}^{|J|}$ such that

$$\begin{bmatrix} 0 & G_I^T H_J \\ H_J^T G_I & 0 \end{bmatrix} \begin{bmatrix} \xi \\ \eta \end{bmatrix} = \cos\theta \begin{bmatrix} G_I^T G_I & 0 \\ 0 & H_J^T H_J \end{bmatrix} \begin{bmatrix} \xi \\ \eta \end{bmatrix}, \tag{3.39}$$

$$\langle g_k, H_J\eta - (\cos\theta) G_I\xi \rangle \geq 0 \quad \textit{for all } k \notin I, \tag{3.40}$$

$$\langle h_\ell, G_I\xi - (\cos\theta) H_J\eta \rangle \geq 0 \quad \textit{for all } \ell \notin J, \tag{3.41}$$

$$\langle \xi, G_I^T G_I \xi \rangle = 1, \ \xi \in \operatorname{int}(\mathbb{R}^{|I|}_+), \tag{3.42}$$

$$\langle \eta, H_J^T H_J \eta \rangle = 1, \ \eta \in \operatorname{int}(\mathbb{R}^{|J|}_+). \tag{3.43}$$

Furthermore, when these equivalent statements hold, the critical angle θ is formed with the critical pair $(u, v) = (G_I\xi, H_J\eta)$.

Proof. We follow similar steps as in [9, Theorem 3], except that now the polyhedral cones P and Q are not necessarily equal. Besides, we do not restrict the attention to proper critical angles. For the sake of completeness, we give a sketch of the proof:
$(a) \Rightarrow (b)$. Let (u, v) be a critical pair of (P, Q) such that $\lambda := \langle u, v \rangle = \cos\theta$. The cone version of Caratheodory's theorem ensures the existence of index sets $I \in \mathcal{I}(G)$, $J \in \mathcal{J}(H)$, and vectors $\xi \in \mathbb{R}^{|I|}$, $\eta \in \mathbb{R}^{|J|}$ with positive components, such that $u = G_I\xi$ and $v = H_J\eta$. The normalization conditions in (3.42)-(3.43) are obtained from the fact that u and v are unit vectors. Criticality of (u, v) leads to the system

$$\begin{cases} H_J\eta - \lambda G_I\xi \in P^*, \\ G_I\xi - \lambda H_J\eta \in Q^*, \end{cases}$$

or, equivalently,

$$\begin{aligned} \langle g_k, H_J\eta - \lambda G_I\xi\rangle &\geq 0 \quad \text{for all } k = 1,\dots,p, \\ \langle h_\ell, G_I\xi - \lambda H_J\eta\rangle &\geq 0 \quad \text{for all } \ell = 1,\dots,q. \end{aligned}$$

This yields (3.40) and (3.41). Furthermore, since

$$\begin{aligned} 0 = \langle u, v - \lambda u\rangle &= \langle \xi, G_I^T H_J\eta - \lambda G_I^T G_I\xi\rangle, \\ 0 = \langle v, u - \lambda v\rangle &= \langle \eta, H_J^T G_I\xi - \lambda H_J^T H_J\eta\rangle, \end{aligned}$$

one gets

$$\begin{aligned} G_I^T H_J\eta - \lambda G_I^T G_I\xi &= 0, \\ H_J^T G_I\xi - \lambda H_J^T H_J\eta &= 0, \end{aligned}$$

which is nothing but (3.39).
$(b) \Rightarrow (a)$. If one sets $(u, v) := (G_I\xi, H_J\eta)$, then one can check that (u, v) is a critical pair of (P, Q) with $\cos\theta = \langle u, v\rangle$. □

The index sets I, J and the vectors ξ, η in Theorem 3.22(b) are not necessarily unique. Anyway, one can write

$$\Gamma(P, Q) = \bigcup_{I\in\mathcal{I}(G)} \bigcup_{J\in\mathcal{J}(H)} \Gamma_{I,J}(P, Q),$$

where $\Gamma_{I,J}(P, Q)$ captures the critical angles produced by (I, J), that is,

$$\Gamma_{I,J}(P, Q) := \{\arccos\langle G_I\xi, H_J\eta\rangle : (\xi, \eta) \text{ as in (3.39)-(3.43)}\}.$$

One refers to (I, J) as a *successful configuration of index sets* if $\Gamma_{I,J}(P, Q)$ is nonempty. For each pair (I, J), we construct $\Gamma_{I,J}(P, Q)$ by using the following algorithm:

- Step 1: Solve the generalized eigenvalue problem $A^{I,J}z = \lambda B^{I,J}z$, where

$$A^{I,J} := \begin{bmatrix} 0 & G_I^T H_J \\ H_J^T G_I & 0 \end{bmatrix}, \; B^{I,J} := \begin{bmatrix} G_I^T G_I & 0 \\ 0 & H_J^T H_J \end{bmatrix}, \; z := \begin{bmatrix} \xi \\ \eta \end{bmatrix}.$$

- Step 2: Declare "acceptable" each eigenvalue that admits an associated eigenvector satisfying the conditions (3.40)-(3.43), where one identifies λ with $\cos\theta$. Take the arccosinus of each acceptable eigenvalue and put it in the set $\Gamma_{I,J}(P, Q)$.

Example 3.23. By way of example, consider the nonnegative orthant $P = \mathbb{R}^n_+$ and the Schur cone

$$Q = \left\{ x \in \mathbb{R}^n : \sum_{i=1}^{k} x_i \geq 0 \text{ for } k \in \{1,\dots,n-1\} \text{ and } x_1 + \ldots + x_n = 0 \right\}.$$

In this case, $G = I_n$ and H is formed with the n-dimensional vectors

$$h_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \; \ldots, \; h_{n-1} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ -1 \end{pmatrix}.$$

Table 3.1 concerns the particular case $n = 5$. It displays the successful configurations (I, J) and the critical angles produced by each one of these configurations. As one can see from Table 3.1, different configurations (I, J) may produce the same critical angle. There are 465 configurations (I, J) in all, but only 12 are successful.

I	J	$\cos\theta$	θ
$\{2,3,4,5\}$	$\{1,2,3,4\}$	$-1/\sqrt{5}$	$0.6476\,\pi$
$\{2,3,4\}$	$\{1,2,3\}$	$-1/2$	$0.6667\,\pi$
$\{2,3\}$	$\{1,2\}$	$-1/\sqrt{3}$	$0.6959\,\pi$
$\{2,4,5\}$	$\{1,2,3,4\}$	$-\sqrt{2}/\sqrt{5}$	$0.7180\,\pi$
$\{3,4,5\}$	$\{1,2,3,4\}$	$-\sqrt{2}/\sqrt{5}$	$0.7180\,\pi$
$\{2\}$	$\{1\}$	$-1/\sqrt{2}$	$0.7500\,\pi$
$\{3,4\}$	$\{1,2,3\}$	$-1/\sqrt{2}$	$0.7500\,\pi$
$\{3,5\}$	$\{1,2,3,4\}$	$-\sqrt{3}/\sqrt{5}$	$0.7820\,\pi$
$\{4,5\}$	$\{1,2,3,4\}$	$-\sqrt{3}/\sqrt{5}$	$0.7820\,\pi$
$\{3\}$	$\{1,2\}$	$-\sqrt{2}/\sqrt{3}$	$0.8041\,\pi$
$\{4\}$	$\{1,2,3\}$	$-\sqrt{3}/2$	$0.8333\,\pi$
$\{5\}$	$\{1,2,3,4\}$	$-2/\sqrt{5}$	$0.8524\,\pi$

Table 3.1: Critical angles between the nonnegative orthant and the Schur cone in $\mathbb{R}^5$.

Remark 3.24. It is not surprising that all the critical angles reported in Table 3.1 are obtuse. This corresponds to a general fact concerning critical angles between nonnegative orthants and Schur cones.

3.7.1 Uniform cardinality estimates for angular spectra

Let $\mathcal{M}_n(p, q)$ denote the collection of all pairs (P, Q) as in (3.37)-(3.38). Theorem 3.22 yield the following uniform cardinality estimate.

Proposition 3.25. *Fix the dimension n. The term*

$$c_n(p, q) := \max_{(P,Q)\in\mathcal{M}_n(p,q)} \operatorname{card}[\Gamma(P, Q)]$$

is majorized by a bivariate polynomial, of degree at most $2n$, in the variables p and q. In particular, $c_n(p, q)$ grows at most polynomially in each argument separately.

Proof. The generalized eigenvalue problem (3.39) has at most $|I| + |J|$ eigenvalues. If one defines

$$C_k^p := \begin{cases} \frac{p!}{k!(p-k)!} & \text{if } k \le p, \\ 0 & \text{if } k > p, \end{cases} \tag{3.44}$$

then Theorem 3.22 shows that

$$\operatorname{card}[\Gamma(P, Q)] \le \sum_{k=1}^{n} \sum_{\ell=1}^{n} C_k^p C_\ell^q (k + \ell)$$

for all $(P, Q) \in \mathcal{M}_n(p, q)$. One gets in this way

$$c_n(p,q) \leq \left(\sum_{k=1}^{n} kC_k^p\right)\left(\sum_{\ell=1}^{n} C_\ell^q\right) + \left(\sum_{\ell=1}^{n} \ell C_\ell^q\right)\left(\sum_{k=1}^{n} C_k^p\right). \tag{3.45}$$

Note that each sum between parentheses defines a polynomial of degree at most n, either in the variable p or in the variable q. The term on the right-hand side of (3.45) is then a bivariate polynomial, of degree at most $2n$, with respect to (p, q). □

The term on the right-hand side of (3.45) is a rather coarse upper bound for $c_n(p, q)$. Lower bounds for $c_n(p, q)$ can be obtained experimentally by counting critical angles in randomly generated pairs of polyhedral cones. For each choice of (p, q), the lower bound displayed in Table 3.2 was obtained by working with a sample of 10^4 randomly generated pairs of polyhedral cones. The columns of G and the columns of H are random vectors uniformly distributed on the nonnegative orthant intersected with the unit sphere of $\mathbb{R}^n$.

	$p = n = 3$	$p = n = 4$	$p = n = 5$	$p = n = 6$
q=1	7	15	31	63
q=2	11	31	53	99
q=3	13	37	77	151
q=4	–	41	87	185
q=5	–	–	105	217
q=6	–	–	–	247

Table 3.2: Lower bounds for $c_n(p, q)$.

For instance, when $q = p = n = 4$, one can find a pair (P, Q) with as much as 41 critical angles. Experimentally, we found the pair (P, Q) described by the matrices

$$G = \begin{bmatrix} 0.547 & 0.347 & 0.255 & 0.425 \\ 0.109 & 0.756 & 0.807 & 0.273 \\ 0.815 & 0.550 & 0.490 & 0.339 \\ 0.152 & 0.066 & 0.202 & 0.793 \end{bmatrix}, \quad H = \begin{bmatrix} 0.408 & 0.843 & 0.041 & 0.455 \\ 0.675 & 0.187 & 0.556 & 0.119 \\ 0.601 & 0.481 & 0.797 & 0.485 \\ 0.120 & 0.148 & 0.228 & 0.737 \end{bmatrix}.$$

3.7.2 A polyhedral cone versus a ray

We now search for critical angles between a polyhedral cone P and a ray

$$\mathbb{R}_+ v := \{tv : t \geq 0\}$$

generated by some unit vector $v \in \mathbb{R}^n$. By adapting Theorem 3.22 to this special setting, one gets the following result.

Proposition 3.26. *Let P be as in (3.37)-(3.38) and v be a unit vector in $\mathbb{R}^n$. Then, $\theta \in \Gamma(P, \mathbb{R}_+ v)$ if and only if there are an index set $I \in \mathcal{I}(G)$ and a vector $\xi \in \mathbb{R}^{|I|}$ such that*

$$\begin{aligned} &(\cos\theta)\xi = (G_I^T G_I)^{-1} G_I^T v, \\ &\langle g_k, v - (\cos\theta) G_I \xi \rangle \geq 0 \ \text{ for all } \ k \notin I, \\ &\langle \xi, G_I^T G_I \xi \rangle = 1, \ \xi \in \operatorname{int}(\mathbb{R}_+^{|I|}) \end{aligned} \tag{3.46}$$

Furthermore, the critical pairs of $(P, \mathbb{R}_+ v)$ *are exactly those of the form* $(G_I \xi, v)$*, with* I *and* ξ *as above.*

Proof. The conditions (3.41) and (3.43) are here superfluous. The generalized eigenvalue problem (3.39) becomes

$$\begin{bmatrix} 0 & G_I^T v \\ v^T G_I & 0 \end{bmatrix} \begin{bmatrix} \xi \\ 1 \end{bmatrix} = \cos\theta \begin{bmatrix} G_I^T G_I & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \xi \\ 1 \end{bmatrix},$$

but this is equivalent to (3.46). □

The next corollary shows that, for each $I \in \mathcal{I}(G)$, the set $\Gamma_I(P, \mathbb{R}_+ v)$ is empty or a singleton. The proof of such result is omitted, because it follows straightforwardly from Proposition 3.26.

Corollary 3.27. *Let* P *be as in* (3.37)-(3.38) *and* v *be a unit vector in* $\mathbb{R}^n$*. For each* $I \in \mathcal{I}(G)$*, consider the condition*

$$\langle g_k, v - v^I \rangle \geq 0 \quad \text{for all } k \notin I, \tag{3.47}$$

where $v^I := G_I (G_I^T G_I)^{-1} G_I^T v$ *is the orthogonal projection of* v *onto* $\mathrm{Im} G_I$*. Then*

$$\Gamma_I(P, \mathbb{R}_+ v) = \begin{cases} \{\pi/2\} & \text{if } G_I^T v = 0 \text{ and (3.47) holds,} \\ \{\arccos \|v^I\|\} & \text{if } (G_I^T G_I)^{-1} G_I^T v \in \mathrm{int}(\mathbb{R}_+^{|I|}) \text{ and (3.47) holds,} \\ \{\pi - \arccos \|v^I\|\} & \text{if } -(G_I^T G_I)^{-1} G_I^T v \in \mathrm{int}(\mathbb{R}_+^{|I|}) \text{ and (3.47) holds,} \\ \emptyset & \text{otherwise.} \end{cases}$$

Since each index set $I \in \mathcal{I}(G)$ produces at most one critical angle, one gets the upper bound

$$\mathrm{card}[\Gamma(P, \mathbb{R}_+ v)] \leq \sum_{k=1}^{n} C_k^p, \tag{3.48}$$

where C_k^p is given by (3.44). The sum on the right-hand side of (3.48) is a polynomial, of degree at most n, in the variable p. By combining the bound (3.48) and the row $q = 1$ of Table 3.2, one sees that

$$c_n(n, 1) = 2^n - 1,$$

at least when $n \in \{3, 4, 5, 6\}$. In practice, the cardinality of $\Gamma(P, \mathbb{R}_+ v)$ is highly dependent on the orientation of v with respect to the cone P. Consider the random variable

$$\mathbf{x} := \mathrm{card}[\Gamma(P, \mathbb{R}_+ v)],$$

where v and the columns of G are random vectors uniformly distributed on the unit sphere of $\mathbb{R}^n$. Table 3.3 displays the conditional expectations of $\mathbf{x}$ with respect to different orientations of v. Figures are rounded to two decimal places.

	$n=p=3$	$n=p=4$	$n=p=5$	$n=p=6$
$E[\mathbf{x}\mid v\in P]$	4.66	6.46	8.06	9.24
$E[\mathbf{x}\mid v\in -P]$	1	1	1	1
$E[\mathbf{x}\mid v\notin P\cup -P]$	1.40	1.64	1.78	1.86

Table 3.3: Conditional expectations of $\mathbf{x}$ with respect to different orientations of v.

Appendix[4]

Action of invertible linear transformations on angular spectra

What happens with $\Theta(P,Q)$ if the cones $P,Q\in\mathcal{C}(\mathbb{X})$ are subject to a common invertible linear transformation? Consistently with geometric intuition, one has

$$\begin{aligned}\Gamma\left(A(P),A(Q)\right)&=\Gamma(P,Q),\\ \Theta\left(A(P),A(Q)\right)&=\Theta(P,Q)\end{aligned}$$

for any orthogonal linear map $A:\mathbb{X}\to\mathbb{X}$. The above formulas are a particular instance of the next proposition. In what follows, $\mathcal{L}(\mathbb{X})$ denotes the vector space of linear endomorphisms on $\mathbb{X}$ and $\mathbb{GL}(\mathbb{X})$ is the set of invertible linear endomorphisms on $\mathbb{X}$. The superscript "T" stands for transposition or adjunction.

Proposition 3.28. *Let $P,Q\in\mathcal{C}(\mathbb{X})$. Let $A,B\in\mathbb{GL}(\mathbb{X})$ be such that $A^TA=B^TB$. Then*

$$\begin{aligned}\Gamma\left(A(P),A(Q)\right)&=\Gamma\left(B(P),B(Q)\right),\\ \Theta\left(A(P),A(Q)\right)&=\Theta\left(B(P),B(Q)\right).\end{aligned}$$

Proof. Let $\theta\in\Gamma\left(A(P),A(Q)\right)$ and $\lambda:=\cos\theta$. There are vectors $\xi\in P$ and $\eta\in Q$ such that

$$\begin{cases}\lambda=\langle A\xi,A\eta\rangle,\\ \|A\xi\|=1,\ \|A\eta\|=1,\\ A\eta-\lambda A\xi\in[A(P)]^*,\\ A\xi-\lambda A\eta\in[A(Q)]^*.\end{cases}\tag{3.49}$$

The above system can be written in the equivalent form

$$\begin{cases}\lambda=\langle\xi,A^TA\eta\rangle,\\ \langle\xi,A^TA\xi\rangle=1,\ \langle\eta,A^TA\eta\rangle=1,\\ A^TA\eta-\lambda A^TA\xi\in P^*,\\ A^TA\xi-\lambda A^TA\eta\in Q^*.\end{cases}$$

But A^TA can be changed by B^TB. Hence, (3.49) can be written with B in the place of A. This proves that $\theta\in\Gamma\left(B(P),B(Q)\right)$. □

[4]This material is not included in the paper submitted for publication

With the help of a suitable non-orthogonal transformation $A \in \mathbb{GL}(\mathbb{X})$, it is possible to increase (cf. Theorem 3.30) or decrease (cf. Theorem 3.32) arbitrarily the maximal angle of a given pair (P, Q). Before proving this fact, we state a useful lemma.

Lemma 3.29. *Let $(u, v) \in \Omega_{\mathbb{X}}$ and $\vartheta \in]0, \pi[$. Then*

$$\|Bu\| = 1, \quad \|Bv\| = 1, \quad \langle Bu, Bv\rangle = \cos\vartheta \tag{3.50}$$

for some self-adjoint positive definite $B \in \mathcal{L}(\mathbb{X})$.

Proof. Let $n := \dim\mathbb{X}$. The subspace $L := \mathrm{span}\{u, v\}$ has dimension 2. Let $\lambda := \langle u, v\rangle$ and

$$e_1 := (1-\lambda^2)^{-1/2}(u - \lambda v), \; e_2 := v.$$

Let $\{e_3, \ldots, e_n\}$ be an orthonormal basis of the orthogonal complement of L. Let $G \in \mathcal{L}(\mathbb{X})$ be given by

$$Gx := \langle a, x\rangle e_1 + \langle b, x\rangle e_2 + \sum_{k=3}^{n} \langle e_k, x\rangle e_k,$$

where $a, b \in L$ are not collinear. Such a map G is clearly invertible. The particular choice

$$\begin{aligned} a &= (1-\lambda^2)^{-1/2}(\sin\theta)e_1, \\ b &= (1-\lambda^2)^{-1/2}(\cos\theta - \lambda)e_1 + e_2 \end{aligned}$$

yields $Gu = (\sin\vartheta)e_1 + (\cos\vartheta)e_2$, $Gv = e_2$, and $\langle Gu, Gv\rangle = \cos\vartheta$. Since G satisfies the conditions stipulated in (3.50), so does the self-adjoint positive definite linear map $B := \left(G^T G\right)^{1/2}$. □

Theorem 3.30. *Let $P, Q \in \mathcal{C}(\mathbb{X})$ and θ be such that $0 < \Theta(P, Q) < \theta < \pi$. Then*

$$\Theta\left(A(P), A(Q)\right) = \theta \tag{3.51}$$

for some $A \in \mathbb{GL}(\mathbb{X})$.

Proof. Let $u \in P \cap S_{\mathbb{X}}$ and $v \in Q \cap S_{\mathbb{X}}$ be non-collinear. Take any $\vartheta \in]\theta, \pi[$ and construct B as in Lemma 3.29. There exists a continuous function $\mathbf{A} : [0, 1] \to \mathcal{L}(\mathbb{X})$ such that $\mathbf{A}(0) = I_{\mathbb{X}}$, $\mathbf{A}(1) = B$, and $\mathbf{A}(t) \in \mathbb{GL}(\mathbb{X})$ for all $t \in [0, 1]$. To see this, just consider the obvious choice

$$\mathbf{A}(t) := (1-t)I_{\mathbb{X}} + tB. \tag{3.52}$$

The real-valued function

$$t \in [0, 1] \mapsto f(t) := \Theta\left(\mathbf{A}(t)(P), \mathbf{A}(t)(Q)\right) \tag{3.53}$$

is continuous, because $\Theta : \mathcal{C}^2(\mathbb{X}) \to \mathbb{R}$ is continuous and $t \mapsto (\mathbf{A}(t)(P), \mathbf{A}(t)(Q))$ is continuous as function from $[0, 1]$ to $\mathcal{C}^2(\mathbb{X})$. On the other hand, $f(0) = \Theta(P, Q)$ and

$$f(1) = \Theta\left(B(P), B(Q)\right) \geq \arccos\langle Bu, Bv\rangle = \vartheta.$$

The Intermediate Value Theorem ensures the existence of $t_* \in [0, 1]$ such that $f(t_*) = \theta$. Hence, (3.51) holds with $A = \mathbf{A}(t_*)$. □

We now explain how to construct a non-orthogonal transformation $A \in \mathbb{GL}(\mathbb{X})$ that decreases arbitrarily the maximal angle of a pair (P, Q). Again, a preliminary lemma is in order.

Lemma 3.31. *Let $K \in \mathcal{C}(\mathbb{X})$ be such that $0 < \theta(K) < \pi$. Then, for all $\psi \in]0, \pi[$, there exists $A \in \mathbb{GL}(\mathbb{X})$ such that $\theta_{\max}(A(K)) = \psi$.*

Proof. When $\psi = \theta_{\max}(K)$, one takes $A = I_{\mathbb{X}}$. When $\psi > \theta_{\max}(K)$, one applies Theorem 3.30 with $P = Q = K$. Suppose now that $\psi < \theta_{\max}(K)$. Pick any ϑ such that $0 < \vartheta < \min\{\psi, \pi/2\}$. Let $n := \dim\mathbb{X}$. The assumption $\theta_{\max}(K) < \pi$ implies that K^* contains a basis $\{f_1, \ldots, f_n\}$ of $\mathbb{X}$. Hence,

$$K \subseteq S := \{x \in \mathbb{X} : \langle f_1, x\rangle \geq 0, \ldots, \langle f_n, x\rangle \geq 0\}.$$

Note that S is a simplicial cone in $\mathbb{X}$. It can be represented in the form

$$S = G\left(\mathbb{R}^n_+\right) = \left\{\sum_{i=1}^n \xi_i g_i : \xi \in \mathbb{R}^n_+\right\},$$

where $\{g_1, \ldots, g_n\}$ is another basis of $\mathbb{X}$ and $G : \mathbb{R}^n \to \mathbb{X}$ is given by $G\xi := \sum_{i=1}^n \xi_i g_i$. Consider a matrix $C = [c_1, \ldots, c_n]$ of order n whose columns are linearly independent unit vectors satisfying

$$\langle c_i, c_j\rangle \geq \cos\vartheta \quad \text{for all } i \neq j.$$

Thanks to [6, Proposition 6.2], one has $\theta_{\max}\left(C(\mathbb{R}^n_+)\right) \leq \vartheta$. Let $H : \mathbb{X} \to \mathbb{R}^n$ be the linear map given by $Hx := CG^{-1}x$. Since

$$H(K) \subseteq H(S) = C\left(\mathbb{R}^n_+\right),$$

one gets $\theta_{\max}(H(K)) \leq \vartheta$. Proposition 3.28 yields $\theta_{\max}(B(K)) \leq \vartheta$ with $B := (H^T H)^{1/2}$. Next, one considers $\mathbf{A}$ given by (3.52) and applies the Intermediate Value Theorem to $f(t) := \theta_{\max}(\mathbf{A}(t)(K))$. Since $f(0) = \theta_{\max}(K)$ and $f(1) \leq \vartheta$, there exists $t_* \in [0, 1]$ such that $\theta_{\max}(\mathbf{A}(t_*)(K)) = \psi$. □

Theorem 3.32. *Let $P, Q \in \mathcal{C}(\mathbb{X})$ and θ be such that $0 < \theta < \Theta(P, Q) < \pi$. In addition, suppose that P and Q are pointed separately. Then there exists $A \in \mathbb{GL}(\mathbb{X})$ such that $\Theta(A(P), A(Q)) = \theta$.*

Proof. Pick any $\vartheta \in]0, \theta[$. Under the hypotheses of the theorem, the convex cone $K := P + Q$ is closed, pointed, and contains a pair of non-collinear vectors. By applying Lemma 3.31 one obtains

$$\Theta(G(P), G(Q)) \leq \theta_{\max}(G(K)) = \vartheta \tag{3.54}$$

for a suitable $G \in \mathbb{GL}(\mathbb{X})$. Of course, one can write (3.54) with $B := \left(G^T G\right)^{1/2}$ instead of G. It remains now to apply the Intermediate Value Theorem to the function (3.53), with $\mathbf{A}$ given by (3.52). □

Bibliography

[1] J.-P. Aubin and H. Frankowska. *Set-Valued Analysis*. Birkhäuser, Boston, 1990.

[2] E. Beutner. On the closedness of the sum of closed convex cones in reflexive Banach spaces. *J. Convex Anal.*, 14 (2007), 99–102.

[3] A. Björck and G.H. Golub. Numerical methods for computing angles between linear subspaces. *Math. Comp.*, 27 (1973), 579–594.

[4] F. Goldberg and N. Shaked-Monderer. On the maximal angle between copositive matrices. July 2013, submitted. Temporarily available at http://arxiv.org/pdf/1307.7519.pdf.

[5] D. Gourion and A. Seeger. Critical angles in polyhedral convex cones: numerical and statistical considerations. *Math. Program.*, 123 (2010), 173-198.

[6] A. Iusem and A. Seeger. On pairs of vectors achieving the maximal angle of a convex cone. *Math. Program.*, 104 (2005), 501–523.

[7] A. Iusem and A. Seeger. Angular analysis of two classes of non-polyhedral convex cones: the point of view of optimization theory. *Comput. Applied Math.*, 26 (2007), 191–214.

[8] A. Iusem and A. Seeger. On convex cones with infinitely many critical angles. *Optimization*, 56 (2007), 115–128.

[9] A. Iusem and A. Seeger. Antipodal pairs, critical pairs, and Nash angular equilibria in convex cones. *Optim. Meth. Software*, 23 (2008), 73–93.

[10] A. Iusem and A. Seeger. Normality and modulability indices. II. Convex cones in Hilbert spaces. *J. Math. Anal. Appl.*, 338 (2008), 392–406.

[11] A. Iusem and A. Seeger. Searching for critical angles in a convex cone. *Math. Program.*, 120 (2009), 3–25.

[12] C. Meyer. *Matrix Analysis and Applied Linear Algebra*. SIAM Publications, Philadelphia, 2000.

[13] J.M. Miao and A. Ben-Israel. On principal angles between subspaces in R^n. *Linear Algebra Appl.*, 171 (1992), 81–98.

[14] D.G. Obert. The angle between two cones. *Linear Algebra Appl.*, 144 (1991), 63–70.

[15] J. Peña and J. Renegar. Computing approximate solutions for convex conic systems of constraints. *Math. Program.*, 87 (2000), Ser. A, 351–383.

[16] S.N. Roy. A note on critical angles between two flats in hyperspace with certain statistical applications. *Sankhya* 8, (1947), 177–194.

[17] A. Seeger. Lipschitz and Hölder continuity results for some functions of cones. *Positivity*, online September 2013, DOI: 10.1007/s11117-013-0258-0.

[18] A. Seeger and M. Torki. On eigenvalues induced by a cone constraint. *Linear Algebra Appl.*, 372 (2003), 181–206.

[19] M. Tenenhaus. Canonical analysis of two convex polyhedral cones and applications. *Psychometrika* 53 (1988), 503–524.

Chapter 4

Critical angles between two convex cones II. Special cases

Alberto Seeger[1] and David Sossa[2]

Abstract. The concept of critical angle between two linear subspaces has applications in statistics, numerical linear algebra, and other areas. Such concept has been abundantly studied in the literature. Part I of this work is an attempt to build up a theory of critical angles for a pair of closed convex cones. Part II focusses on the practical computation of the maximal angle between specially structured cones.

Mathematics Subject Classification: 15A18, 15A48, 52A40, 90C26, 90C33.
Key words: Maximal angle, critical angle, convex cones, topheavy cones, ellipsoidal cones, cones of matrices.

4.1 Introduction

Let $(\mathbb{X}, \langle\cdot,\cdot\rangle)$ be a Euclidian space of dimension at least two and let $\mathcal{C}(\mathbb{X})$ be the set of nontrivial closed convex cones in $\mathbb{X}$. That a closed convex cone is nontrivial means that it is different from the zero cone and different from the whole space. The maximal angle of a pair (P, Q) of nontrivial closed convex cones in $\mathbb{X}$ is defined by

$$\Theta(P,Q) \;:=\; \max_{u\in P\cap S_{\mathbb{X}}\,,\, v\in Q\cap S_{\mathbb{X}}} \arccos\langle u, v\rangle, \tag{4.1}$$

[1]Université d'Avignon, Département de Mathématiques, 33 rue Louis Pasteur, 84000 Avignon, France (alberto.seeger@univ-avignon.fr)

[2]Departamento de Ingeniería Matemática, Centro de Modelamiento Matemático (CNRS UMI 2807), FCFM, Universidad de Chile, Blanco Encalada 2120, Santiago, Chile (dsossa@dim.uchile.cl). This author is supported by CONICYT, Chile.

where $S_{\mathbb{X}}$ stands for the unit sphere of $\mathbb{X}$. A pair $(u, v) \in \mathbb{X}^2$ solving the angle maximization problem (4.1) is called an *antipodal pair* of (P, Q). Antipodal pairs always exist, but they are not unique in general. A necessary condition for (u, v) to be an antipodal pair of (P, Q) is that

$$\begin{cases} u \in P \cap S_{\mathbb{X}}, \\ v \in Q \cap S_{\mathbb{X}}, \\ v - \langle u, v\rangle u \in P^*, \\ u - \langle u, v\rangle v \in Q^*, \end{cases} \tag{4.2}$$

where P^* and Q^* are the positive dual cones of P and Q, respectively. If the system (4.2) holds, then (u, v) is called a *critical pair* of (P, Q) and $\arccos\langle u, v\rangle$ is called a *critical angle* of (P, Q). The adjective *proper* is added to a critical pair (u, v) and the corresponding critical angle if u and v are not collinear. One refers to the set

$$\Gamma(P, Q) := \{\arccos\langle u, v\rangle : (u, v) \text{ satisfies (4.2)}\}$$

as the *angular spectrum* of (P, Q). Angular spectra have usually a finite cardinality, but not always.

In a similar way one can treat the angle minimization problem

$$\Psi(P, Q) := \min_{u \in P \cap S_{\mathbb{X}},\, v \in Q \cap S_{\mathbb{X}}} \arccos\langle u, v\rangle. \tag{4.3}$$

Angle minimization problems like (4.3) arise in a number of applications, for instance in the theory of exponential dichotomies for linear ODEs (cf. [5]) and in regression analysis of ordinal data (cf. [9]). One readily sees that

$$\begin{aligned} \cos[\Psi(P, Q)] &= -\cos[\Theta(P, -Q)], \\ \Psi(P, Q) &= \pi - \Theta(P, -Q). \end{aligned} \tag{4.4}$$

So, there is no loss of generality in focussing the attention just on angle maximization.

Part I of this work (cf. [8]) establishes various geometric and analytic results concerning antipodality, criticality, and angular spectra. The present paper focusses on the computation of the maximal angle between specially structured cones. The organization of the paper is as follows.

- Section 4.2 discusses the case in which P and Q are revolution cones. We give explicit formulas for computing all the critical angles.
- Section 4.3 discusses the case in which P and Q are topheavy cones. The class of topheavy cones is quite large and include in particular the ℓ^p- cones and the ellipsoidal cones.
- Section 4.4 concerns the computation of the maximal angle between two cones of matrices. A large portion of this section is devoted to a difficult question arising in numerical linear algebra: how large can be the angle between a positive semidefinite symmetric matrix and a symmetric matrix that is nonnegative entrywise?

4.1.1 Preliminary material

A critical pair of (P, Q) may not solve the angle maximization problem (4.1). For instance, if we take $P = \mathbb{R}_+ u$ and $Q = P^*$, with $u \in S_{\mathbb{X}}$, then (u, u) is a critical pair of (P, Q) but it does not solve the problem (4.1) (cf. Proposition 4.2). However, each component of a critical pair is a solution to a certain optimization problem. The details are explained below.

Proposition 4.1. *Let $P, Q \in \mathcal{C}(\mathbb{X})$. Then (u, v) is a critical pair of (P, Q) if and only if*

$$\begin{cases} u \ \textit{minimizes} \ \langle \cdot, v - \langle u, v\rangle u\rangle \ \textit{on} \ P \cap S_{\mathbb{X}}, \\ v \ \textit{minimizes} \ \langle u - \langle u, v\rangle v, \cdot \rangle \ \textit{on} \ Q \cap S_{\mathbb{X}}. \end{cases} \tag{4.5}$$

Proof. The proof is immediate. The key observation is that u is orthogonal to $v - \langle u, v\rangle u$ and that v is orthogonal to $u - \langle u, v\rangle v$. □

There are many alternative characterizations of criticality. The characterization (4.5) will be used later in a number of occasions. Beware that (4.5) is a weaker than

$$\begin{cases} u \ \text{minimizes} \ \langle \cdot, v\rangle \ \text{on} \ P \cap S_{\mathbb{X}}, \\ v \ \text{minimizes} \ \langle u, \cdot \rangle \ \text{on} \ Q \cap S_{\mathbb{X}}. \end{cases} \tag{4.6}$$

A pair (u, v) as in (4.6) is said to be a *Nash antipodal pair* of (P, Q). Nash antipodality is a property that lies between criticality and antipodality.

The following easy result is recorded just for convenience. It concerns the maximal angle between a convex cone and its dual.

Proposition 4.2. *Let $K \in \mathcal{C}(\mathbb{X})$. Then $\Theta(K, K^*) = \pi/2$.*

Proof. Since $\langle u, v\rangle \geq 0$ for all $u \in K$ and $v \in K^*$, it is clear that

$$\Theta(K, K^*) \leq \pi/2. \tag{4.7}$$

Since $K \cup -K^*$ is not the whole space $\mathbb{X}$, there exists a nonzero vector $z \notin K \cup -K^*$. Let $\Pi_K(z)$ denote the projection of z onto K. Moreau's orthogonal decomposition theorem implies that $\Pi_K(z)$ and $\Pi_{K^*}(-z)$ are nonzero orthogonal vectors. Hence,

$$u := \frac{\Pi_K(z)}{\|\Pi_K(z)\|} \in K, \quad v := \frac{\Pi^*_K(-z)}{\|\Pi^*_K(-z)\|} \in K^*$$

are orthogonal unit vectors. This proves that (4.7) is in fact an equality. □

4.2 Critical angles in a pair of revolution cones

Revolutions cones, also called circular cones, are amongst the simplest and most common non-polyhedral convex cones used in mathematics. By definition, a revolution cone in $\mathbb{X}$ is a closed convex cone of the form

$$\mathrm{Rev}(\phi, b) := \{x \in \mathbb{X} : \langle b, x\rangle \geq \|x\| \cos\phi\},$$

where $b \in S_{\mathbb{X}}$ defines the revolution axis and $\phi \in [0, \pi/2]$ corresponds to the half-aperture angle. The next theorem shows that a pair of revolution cones has at most three critical angles. It provides also explicit formulas for computing each critical angle.

Theorem 4.3. *Let* $P = \mathrm{Rev}(\phi_1, b_1)$ *and* $Q = \mathrm{Rev}(\phi_2, b_2)$*, with* $b_1, b_2 \in S_{\mathbb{X}}$ *and* $\phi_1, \phi_2 \in [0, \pi/2]$*. Then*

$$\Gamma(P, Q) = \begin{cases} \{0, \alpha_1, \pi\} & \text{if } \alpha_1 \geq 0,\ \alpha_2 \geq \pi, \\ \{0, \alpha_1, \alpha_2\} & \text{if } \alpha_1 \geq 0,\ \alpha_2 < \pi, \\ \{\pi\} & \text{if } \alpha_1 < 0,\ \alpha_2 \geq \pi, \\ \{\alpha_2\} & \text{if } \alpha_1 < 0,\ \alpha_2 < \pi, \end{cases} \tag{4.8}$$

where

$$\begin{aligned} \alpha_1 &:= \phi_1 + \phi_2 - \arccos\langle b_1, b_2\rangle, \\ \alpha_2 &:= \phi_1 + \phi_2 + \arccos\langle b_1, b_2\rangle. \end{aligned}$$

Proof. If $\mathbb{X}$ is a two dimensional space, then (4.8) is obtained by arguments of planar geometry. Suppose that $\mathbb{X}$ is of dimension at least three. The improper critical angles of (P, Q) are easy to identify. Indeed,

$$\begin{aligned} 0 \in \Gamma(P, Q) &\Leftrightarrow P \cap Q \neq \{0\} \Leftrightarrow \alpha_1 \geq 0, \\ \pi \in \Gamma(P, Q) &\Leftrightarrow P \cap -Q \neq \{0\} \Leftrightarrow \alpha_2 \geq \pi. \end{aligned}$$

So, one just needs to detect the proper critical angles of (P, Q). We claim that

$$\begin{cases} \text{if } (u, v) \text{ is a proper critical pair of } (P, Q), \\ \text{then } L := \mathrm{span}\{u, v\} \text{ contains } b_1 \text{ and } b_2. \end{cases} \tag{4.9}$$

This claim will be shown in a moment. As a consequence of (4.9), one gets the following planar reduction principle:

$$\begin{cases} (u, v) \text{ is a proper critical pair of } (P, Q) \text{ if and only if} \\ (u, v) \text{ is a proper critical pair of } (P \cap L, Q \cap L), \end{cases}$$

where

$$\begin{aligned} P \cap L &:= \{x \in L : \langle b_1, x\rangle \geq \|x\| \cos\phi_1\}, \\ Q \cap L &:= \{x \in L : \langle b_2, x\rangle \geq \|x\| \cos\phi_2\}, \end{aligned}$$

are viewed as revolution cones in the two dimensional space L. In other words, one is back to a planar setting. We now prove (4.9). Consider a proper critical pair (u, v) of (P, Q) and

set $\lambda := \langle u, v\rangle$. We distinguish between two cases:
Case 1: $\phi_1, \phi_2 > 0$. This case is the most interesting one. By combining Proposition 4.1 and the boundary principle for proper critical pairs (cf. [8, Theorem 2.3]), one knows that that u and v solve

$$\begin{cases} \text{minimize } \langle x, v - \lambda u\rangle \\ \langle b_1, x\rangle = \cos\phi_1 \\ \|x\|^2 = 1 \end{cases} \tag{4.10}$$

and

$$\begin{cases} \text{minimize } \langle u - \lambda v, y\rangle \\ \langle b_2, y\rangle = \cos\phi_2 \\ \|y\|^2 = 1, \end{cases} \tag{4.11}$$

respectively. Hence, there exist Lagrange multipliers μ_1, γ_1 for the problem (4.10) and Lagrange multipliers μ_2, γ_2 for the problem (4.11), such that

$$v - \lambda u = \gamma_1 b_1 + \mu_1 u, \tag{4.12}$$
$$u - \lambda v = \gamma_2 b_2 + \mu_2 v. \tag{4.13}$$

The multiplication of (4.12) by u and (4.13) by v yields

$$\mu_1 + \gamma_1 \cos\phi_1 = 0, \quad \mu_2 + \gamma_2 \cos\phi_2 = 0,$$

respectively. Hence,

$$\begin{aligned} (\gamma_1 \cos\phi_1 - \lambda)u + v &= \gamma_1 b_1, \\ u + (\gamma_2 \cos\phi_2 - \lambda)v &= \gamma_2 b_2. \end{aligned} \tag{4.14}$$

Observe that $\gamma_1 \neq 0$ and $\gamma_2 \neq 0$, because (u, v) is proper. This proves that $b_1, b_2 \in L$.

Case 2: $\phi_1 = 0, \phi_2 > 0$. In this case, $P = \mathbb{R}_+ b_1$. Hence, $u = b_1$ and $b_1 \in L$. That $b_2 \in L$ follows from (4.14) and the fact that $\gamma_2 \neq 0$. □

Corollary 4.4. *Let P, Q be two revolution cones as in Theorem 4.3. Then*

$$\Theta(P, Q) := \min\{\pi, \arccos\langle b_1, b_2\rangle + \phi_1 + \phi_2\}, \tag{4.15}$$
$$\Psi(P, Q) := \max\{0, \arccos\langle b_1, b_2\rangle - \phi_1 - \phi_2\}. \tag{4.16}$$

Proof. Formula (4.15) is a direct consequence of Theorem 4.3. Formula (4.16) is obtained by combining (4.4) and (4.15). □

4.3 Maximal angle between two topheavy cones

A topheavy cone in $\mathbb{R}^{n+1}$ is a closed convex cone of the form

$$\mathrm{epi} f := \{(\xi, t) \in \mathbb{R}^{n+1} : f(\xi) \leq t\},$$

where f is a norm on $\mathbb{R}^n$. Topheavy cones are pointed, have nonempty interior, and enjoy some remarkable properties that make them particularly appealing in many applications.

Topheavy cones have been studied under various points of view in the literature, see for instance [1, 4, 7]. The next proposition explains how to compute the maximal angle between two topheavy cones.

Proposition 4.5. *Let f and g be norms on $\mathbb{R}^n$. Then $\cos[\Theta(\mathrm{epi}f, \mathrm{epi}\, g)]$ is equal to the optimal value of the minimization problem*

$$\begin{cases} \text{minimize } \xi \odot \eta \\ \|\xi\|^2 + [f(\xi)]^2 = 1 \\ \|\eta\|^2 + [g(\eta)]^2 = 1, \end{cases} \tag{4.17}$$

where $\odot$ stands for the "product" operation given by

$$\xi \odot \eta := \langle \xi, \eta \rangle + [1 - \|\xi\|^2]^{1/2}[1 - \|\eta\|^2]^{1/2}.$$

Proof. The cones $\mathrm{epi}f$ and $\mathrm{epi}\, g$ have nonempty interior. The first and second components of an antipodal pair must be sough on the boundary of $\mathrm{epi}f$ and on the boundary of $\mathrm{epi}\, g$, respectively. Hence, $\cos[\Theta(\mathrm{epi}f, \mathrm{epi}\, g)]$ is equal to the optimal value of the minimization problem

$$\begin{cases} \text{minimize } \langle \xi, \eta \rangle + ts \\ f(\xi) = t, \\ g(\eta) = s, \\ \|\xi\|^2 + t^2 = 1, \\ \|\eta\|^2 + s^2 = 1. \end{cases}$$

It suffices now to get rid of the variables t and s. □

In order to derive an explicit solution to the problem (4.17), one needs of course a bit more information on the norms f and g. The following definition proves to be useful.

Definition 4.6. *Two norms f, g on $\mathbb{R}^n$ are* lower correlated *if the minimization problems*

$$\alpha_f := \min\{f(x) : \|x\| = 1\}, \tag{4.18}$$

$$\alpha_g := \min\{g(x) : \|x\| = 1\}, \tag{4.19}$$

have a solution in common.

For instance, a positive multiple of $\|\cdot\|$ is lower correlated to any norm on $\mathbb{R}^n$. Without further ado we state:

Theorem 4.7. *Let f, g be lower correlated norms on $\mathbb{R}^n$. Then*

$$\Theta(\mathrm{epi}f, \mathrm{epi}\, g) = \arccos\left(\frac{\alpha_f}{[1+\alpha_f^2]^{1/2}}\right) + \arccos\left(\frac{\alpha_g}{[1+\alpha_g^2]^{1/2}}\right). \tag{4.20}$$

The above maximal angle is attained for instance with the unit vectors

$$(\xi_0, t_0) := \left(\frac{w}{[1+\alpha_f^2]^{1/2}}, \frac{\alpha_f}{[1+\alpha_f^2]^{1/2}}\right) \in \mathrm{epi}f,$$

$$(\eta_0, s_0) := \left(\frac{-w}{[1+\alpha_g^2]^{1/2}}, \frac{\alpha_g}{[1+\alpha_g^2]^{1/2}}\right) \in \mathrm{epi}g,$$

where w is any vector taken from the set

$$S(f,g) := \left[\operatorname{argmin}_{\|x\|=1} f(x)\right] \bigcap \left[\operatorname{argmin}_{\|x\|=1} g(x)\right].$$

Proof. Let $w \in S(f,g)$. Since w is a solution to (4.18) and $-w$ is a solution to (4.19), the vectors

$$\xi_0 := \frac{w}{[1+\alpha_f^2]^{1/2}}, \quad \eta_0 := \frac{-w}{[1+\alpha_g^2]^{1/2}}$$

solve respectively

$$\begin{aligned} \gamma_f &:= \max\{\|\xi\| : \|\xi\|^2 + [f(\xi)]^2 = 1\}, \\ \gamma_g &:= \max\{\|\eta\| : \|\eta\|^2 + [g(\eta)]^2 = 1\}. \end{aligned}$$

Let (ξ,η) be any pair satisfying the equality constraints in (4.17). Then

$$\begin{aligned} \|\xi\| \le \|\xi_0\| = \gamma_f = (1+\alpha_f^2)^{-1/2}, \\ \|\eta\| \le \|\eta_0\| = \gamma_g = (1+\alpha_g^2)^{-1/2}, \end{aligned}$$

and

$$\left\langle \|\xi\|^{-1}\xi, \|\eta\|^{-1}\eta \right\rangle \ge \left\langle \|\xi_0\|^{-1}\xi_0, \|\eta_0\|^{-1}\eta_0 \right\rangle = -1.$$

Hence,

$$\xi \odot \eta \ \ge\ \xi_0 \odot \eta_0 \ =\ -[1+\alpha_f^2]^{-1/2}[1+\alpha_g^2]^{-1/2} + [1-\gamma_f^2]^{1/2}[1-\gamma_g^2]^{1/2}.$$

In other words, (ξ_0,η_0) solves (4.17) and

$$\cos\left[\Theta(\operatorname{epi} f, \operatorname{epi} g)\right] \ =\ \frac{\alpha_f\,\alpha_g - 1}{[1+\alpha_f^2]^{1/2}[1+\alpha_g^2]^{1/2}}.$$

The last equality is an equivalent way of writing (4.20). □

Remark 4.8. In view of [7, Theorem 5.2], formula (4.20) can be rewritten as

$$\Theta(\operatorname{epi} f, \operatorname{epi} g) = \frac{\theta_{\max}(\operatorname{epi} f) + \theta_{\max}(\operatorname{epi} g)}{2}, \tag{4.21}$$

where $\theta_{\max}(K)$ denotes the maximal angle of $K \in \mathcal{C}(\mathbb{X})$. Formula (4.21) is consistent with geometric intuition, but one must remember that f and g are assumed to be lower correlated. Indeed, formula (4.21) may fail if one drops the lower correlation assumption.

As particular example of topheavy cone, one may consider the ℓ^p-cone

$$L_p := \{(\xi,t) \in \mathbb{R}^{n+1} : \|\xi\|_p \le t\}.$$

Here, $p \in [1,\infty]$ and $\|\cdot\|_p$ stands for the ℓ^p-norm in $\mathbb{R}^n$. Of special interest are the cases $p=1$, $p=2$, and $p=\infty$. One gets

$$\begin{aligned} &\Theta(L_1,L_1) = \pi/2, \quad \Theta(L_2,L_2) = \pi/2, \\ &\Theta(L_1,L_2) = \pi/2, \quad \Theta(L_1,L_\infty) = \pi/2, \\ &\Theta(L_\infty,L_\infty) = 2\arccos\left[1+n\right]^{-1/2}, \\ &\Theta(L_2,L_\infty) \ =\ \arccos\left[1+n\right]^{-1/2} + \ \pi/4. \end{aligned}$$

These formulas are obtained by using the proposition stated below.

Proposition 4.9. *The following statements holds:*

(a) Let $p, q \in [2, \infty]$. Then

$$\Theta(L_p, L_q) = \arccos\left[1 + n^{(p-2)/p}\right]^{-1/2} + \arccos\left[1 + n^{(q-2)/q}\right]^{-1/2}.$$

(b) Let $p \in [1, \infty]$ and $q \in [1, p_*]$, where p_* is given by $p^{-1} + p_*^{-1} = 1$. Then

$$\Theta(L_p, L_q) = \pi/2.$$

Proof. Part (a). If p and q are both in $[2, \infty]$, then the norms $f(x) := \|x\|_p$ and $g(x) := \|x\|_q$ are lower correlated. Indeed,

$$\widehat{\mathbf{1}}_n := \frac{1}{\sqrt{n}}(1, \ldots, 1)^T \in \left[\operatorname{argmin}_{\|x\|=1}\|x\|_p\right] \bigcap \left[\operatorname{argmin}_{\|x\|=1}\|x\|_q\right].$$

So, it suffices to substitute $\alpha_f = \|\widehat{\mathbf{1}}_n\|_p$ and $\alpha_g = \|\widehat{\mathbf{1}}_n\|_q$ into (4.20).

Part (b). For all $p \in [1, \infty]$, one has the duality formula $L_p^* = L_{p_*}$ (cf. [4, Proposition 3.1]). Hence, Proposition 4.2 yields $\Theta(L_p, L_{p_*}) = \pi/2$. On the other hand, by applying Theorem 4.7 one gets $\Theta(L_1, L_1) = \pi/2$. Hence, for $q \in [1, p_*]$, one obtains

$$\pi/2 = \Theta(L_1, L_1) \leq \Theta(L_p, L_q) \leq \Theta(L_p, L_{p_*}) = \pi/2. \tag{4.22}$$

Of course, in (4.22) one uses the fact that the family $\{L_p\}_{p \geq 1}$ is nondecreasing with respect to set inclusion. □

4.3.1 Maximal angle between two ellipsoidal cones

An ellipsoidal cone in $\mathbb{R}^{n+1}$ is a closed convex cone of the form

$$E_A := \{(\xi, t) \in \mathbb{R}^{n+1} : \sqrt{\langle \xi, A\xi \rangle} \leq t\},$$

where A is a positive definite symmetric matrix of order n. Hence, an ellipsoidal cone is a particular instance of a topheavy cone. It is easy to see that the norms

$$f(x) = \sqrt{\langle x, Ax \rangle}, \quad g(x) = \sqrt{\langle x, Bx \rangle} \tag{4.23}$$

are lower correlated if the eigenspaces

$$\begin{aligned} E_{\min}(A) &:= \{x \in \mathbb{R}^n : Ax = \lambda_1(A)x\} \\ E_{\min}(B) &:= \{x \in \mathbb{R}^n : Bx = \lambda_1(B)x\} \end{aligned}$$

have a nonzero vector in common. Here, $\lambda_1(A)$ stands for the smallest eigenvalue of A.

Proposition 4.10. *Let A, B be positive definite symmetric matrices of order n. Then $\cos[\Theta(E_A, E_B)]$ is equal to the optimal value of the minimization problem*

$$\begin{cases} \text{minimize } \xi \odot \eta \\ \|\xi\|^2 + \langle \xi, A\xi \rangle = 1 \\ \|\eta\|^2 + \langle \eta, B\eta \rangle = 1. \end{cases}$$

Furthermore, if the eigenspaces $E_{\min}(A)$ and $E_{\min}(B)$ have nonzero vector in common, then

$$\Theta(E_A, E_B) = \arccos\left[\frac{\lambda_1(A)}{1+\lambda_1(A)}\right]^{1/2} + \arccos\left[\frac{\lambda_1(B)}{1+\lambda_1(B)}\right]^{1/2}.$$

Proof. It suffices to apply Proposition 4.5 and Theorem 4.7 to the norms mentioned in (4.23). Note that

$$\alpha_f^2 = \min_{\|x\|=1} \langle x, Ax \rangle = \lambda_1(A).$$

Similarly, $\alpha_g^2 = \lambda_1(B)$. □

4.3.2 An ellipsoidal cone versus a nonnegative orthant

The next proposition gives a formula for computing the maximal angle between an ellipsoidal cone and a nonnegative orthant. We use the notation $\mu_{\min}(C)$ to indicate the smallest Pareto eigenvalue of a square matrix C (cf. [6]). From the general theory of Pareto eigenvalues one knows that

$$\mu_{\min}(C) = \min_{\substack{\|\eta\|=1 \\ \eta \geq 0}} \langle \eta, C\eta \rangle \tag{4.24}$$

whenever the matrix C is symmetric. The notation $\eta \geq 0$ indicates that each component of $\eta \in \mathbb{R}^n$ is nonnegative.

Proposition 4.11. *Let A be a positive definite symmetric matrix of order n. Let $C := -(I_n + A)^{-1}$ with I_n denoting the identity matrix of order n. Then*

$$\Theta(E_A, \mathbb{R}^{n+1}_+) = \arccos\left(-\sqrt{-\mu_{\min}(C)}\right).$$

Furthermore, $((\xi, t), (\eta, s))$ is an antipodal pair of $(E_A, \mathbb{R}^{n+1}_+)$ if and only if

$$\begin{cases} \eta \text{ is a solution to (4.24)}, \\ s = 0, \\ \xi = [-\langle \eta, C\eta \rangle]^{-1/2} C\eta, \\ t = 1 + [\mu_{\min}(C)]^{-1} \|C\eta\|^2. \end{cases}$$

Proof. The term $c := \cos[\Theta(E_A, \mathbb{R}^{n+1}_+)]$ corresponds to the optimal value of the minimization problem

$$\begin{cases} \text{minimize } \langle \xi, \eta \rangle + ts \\ \langle \xi, A\xi \rangle^{1/2} = t, \\ \|\xi\|^2 + t^2 = 1, \\ \|\eta\|^2 + s^2 = 1, \\ \eta \geq 0, s \geq 0. \end{cases} \tag{4.25}$$

Clearly, $s = 0$ at the minimum. By getting rid of the variable t, the problem (4.25) is converted into

$$\begin{cases} \text{minimize } \langle \xi, \eta \rangle \\ \|\xi\|^2 + \langle \xi, A\xi \rangle = 1, \\ \|\eta\|^2 = 1,\ \eta \geq 0 \end{cases}$$

The change of variables $\gamma = (I_n + A)^{1/2}\xi$ leads to

$$c = \min_{\substack{\|\eta\|=1 \\ \eta \geq 0}} \min_{\|\gamma\|=1} \langle (I_n + A)^{-1/2}\eta, \gamma \rangle.$$

Since the inner minimization problem is solved by

$$\gamma = -\|(I_n + A)^{-1/2}\eta\|^{-1} (I_n + A)^{-1/2}\eta,$$

one gets

$$\begin{aligned} -c &= \max_{\substack{\|\eta\|=1 \\ \eta \geq 0}} \|(I_n + A)^{-1/2}\eta\| = \left[\max_{\substack{\|\eta\|=1 \\ \eta \geq 0}} \langle \eta, (I_n + A)^{-1}\eta \rangle \right]^{1/2} \\ &= \left[- \min_{\substack{\|\eta\|=1 \\ \eta \geq 0}} \langle \eta, C\eta \rangle \right]^{1/2} = \left[-\mu_{\min}(C) \right]^{1/2}. \end{aligned}$$

This completes the proof of the proposition. □

4.3.3 An ellipsoidal cone versus a ray

The next proposition explains how to compute the maximal angle between an ellipsoidal cone E_A and a ray $\mathbb{R}_+ v$.

Proposition 4.12. *Let A be a positive definite symmetric matrix of order n and $v := (\eta, s)$ be a unit vector in $\mathbb{R}^{n+1}$. Then*

(a) $\cos[\Theta(E_A, \mathbb{R}_+ v)]$ is equal to the optimal value of the nonconvex minimization problem

$$\begin{cases} \textit{minimize } \langle \xi, \eta \rangle + s[1 - \|\xi\|^2]^{1/2} \\ \|\xi\|^2 + \langle \xi, A\xi \rangle = 1. \end{cases}$$

(b) Under the additional assumption $\langle \eta, A^{-1}\eta \rangle^{1/2} > s > 0$, one can write

$$\cos[\Theta(E_A, \mathbb{R}_+ v)] = -s \min\{\|x - b\| : \langle x, Mx \rangle \leq 1\}, \tag{4.26}$$

where $b := (1/s)(I_n + A)^{-1/2}\eta$ and $M := (I_n + A)^{1/2} A^{-1} (I_n + A)^{1/2}$.

Proof. The proof of (a) is as in Proposition 4.5, so we concentrate on (b). For notational convenience, we write

$$\begin{aligned} f(\xi) &:= \langle \xi, A\xi \rangle^{1/2} = \|A^{1/2}\xi\|, \\ F(\xi) &:= \langle \xi, (I_n + A)\xi \rangle^{1/2} = \|(I_n + A)^{1/2}\xi\|. \end{aligned}$$

Note that f and F are norms on $\mathbb{R}^n$. Let $\gamma := \cos[\Theta(E_A, \mathbb{R}_+ v)]$. One has

$$\begin{aligned} \gamma &= \min_{\substack{(\xi,t)\in E_A \\ \|\xi\|^2+t^2=1}} \{\langle \xi, \eta\rangle + ts\} \\ &= \min_{\substack{f(\xi)=t \\ \|\xi\|^2+t^2=1}} \{\langle \xi, \eta\rangle + ts\} \qquad (4.27) \\ &= \min_{F(\xi)=1} \{\langle \xi, \eta\rangle + sf(\xi)\}, \qquad (4.28) \end{aligned}$$

where (4.27) is a consequence of [8, Theorem 2.3]. Since $E_A^* = E_{A^{-1}}$, the condition $\langle \eta, A^{-1}\eta\rangle^{1/2} > s$ amounts to saying that (η, s) does not belong to dual cone of E_A, i.e., there exists a unit vector $(\tilde{\xi}, \tilde{t}) \in E_A$ such that $\langle \tilde{\xi}, \eta\rangle + \tilde{t}s < 0$. Hence, $\gamma < 0$ and, by a positive homogeneity argument, the equality constraint in (4.28) can be written as an inequality constraint. In other words, one has

$$\gamma = \min_{F(\xi)\leq 1} \{\langle \xi, \eta\rangle + sf(\xi)\}. \qquad (4.29)$$

Observe that (4.29) is a convex minimization problem. By using standard rules of convex analysis, one can show that

$$\gamma = -s \min_{f^\circ(z)\leq 1} F^\circ(z - s^{-1}\eta), \qquad (4.30)$$

where

$$\begin{aligned} f^\circ(\mu) &= \langle \mu, A^{-1}\mu\rangle^{1/2} = \|A^{-1/2}\mu\|, \\ F^\circ(\mu) &= \langle \mu, (I_n + A)^{-1}\mu\rangle^{1/2} = \|(I_n + A)^{-1/2}\mu\| \end{aligned}$$

are the polar norms of f and F, respectively. One can view the minimization problem in (4.30) as a dual version of (4.29). In order to complete the proof of (b), it remains to introduce in (4.30) the change of variables $x = (I_n + A)^{-1/2}z$. □

The minimization problem on the right-hand side of (4.26) is about finding the minimal distance from a point to an ellipsoid. The numerical resolution of such a projection problem offers no difficulty.

4.4 Critical angles between two cones of matrices

Let the space $\mathbb{S}^n$ of symmetric matrices of order n be equipped with the trace inner product $\langle A, B\rangle = \mathrm{tr}(AB)$. This section concerns the analysis of critical angles in a pair of convex cones in $\mathbb{S}^n$. We introduce the symbol $\mathcal{O}_n$ to indicate the set of orthogonal matrices of order n. A nonempty set $\mathcal{P}$ in the space $\mathbb{S}^n$ is said to be *orthogonally invariant* if

$$A \in \mathcal{P} \;\Rightarrow\; U^T A U \in \mathcal{P} \;\text{ for all } U \in \mathcal{O}_n.$$

For instance, the SDP cone

$$\mathbb{S}^n_+ := \{A \in \mathbb{S}^n : A \text{ is positive semidefinite}\}$$

is orthogonally invariant.

Proposition 4.13. *Suppose that at least one of the cones $\mathcal{P}, \mathcal{Q} \in \mathcal{C}(\mathbb{S}^n)$ is orthogonally invariant. Let (A, B) be a critical pair of $(\mathcal{P}, \mathcal{Q})$. Then A and B commute, i.e., $AB = BA$.*

Proof. Suppose, for instance, that $\mathcal{P}$ is orthogonally invariant. Write $\lambda := \langle A, B\rangle$. By Proposition 4.1 one knows that A minimizes the linear form $\langle B - \lambda A, \cdot\,\rangle$ on

$$\mathcal{P}^\diamond := \{X \in \mathcal{P} : \|X\| = 1\}.$$

Since $\mathcal{P}^\diamond$ is an orthogonally invariant set, the commutation principle stated in [3, Lemma 4] implies that $A(B - \lambda A) = (B - \lambda A)A$. This leads to $AB = BA$. □

There is a rich literature devoted to the analysis of orthogonally invariant sets. One knows, for instance, that $\mathcal{P} \in \mathcal{C}(\mathbb{S}^n)$ is orthogonally invariant if and only if there exists a permutation invariant cone $P \in \mathcal{C}(\mathbb{R}^n)$ such that

$$\mathcal{P} = \lambda^{-1}(P) := \{A \in \mathbb{S}^n : \lambda(A) \in P\}. \tag{4.31}$$

Here and in the sequel, the notation $\lambda(A)$ stands for the vector of eigenvalues of A arranged in nondecreasing order, i.e.,

$$\lambda_1(A) \leq \ldots \leq \lambda_n(A).$$

The cone P in the representation formula (4.31) is unique and given by

$$P = \{x \in \mathbb{R}^n : \mathrm{Diag}(x) \in \mathcal{P}\},$$

where $\mathrm{Diag}(x)$ is the diagonal matrix whose entries on the diagonal are the components of x. One refers to P as the permutation invariant cone associated to $\mathcal{P}$.

Theorem 4.14. *Let $\mathcal{P}, \mathcal{Q} \in \mathcal{C}(\mathbb{S}^n)$ be orthogonally invariant and $P, Q \in \mathcal{C}(\mathbb{R}^n)$ be the associated permutation invariant cones. Then*

$$\Gamma(\mathcal{P}, \mathcal{Q}) = \Gamma(P, Q), \quad \Theta(\mathcal{P}, \mathcal{Q}) = \Theta(P, Q).$$

Furthermore, the following statements are equivalent:

(a) (A, B) is a critical (respectively, antipodal) pair of $(\mathcal{P}, \mathcal{Q})$,

(b) there exist a critical (respectively, antipodal) pair (u, v) of (P, Q) and a matrix $U \in \mathcal{O}_n$ such that $A = U\mathrm{Diag}(u)U^T$ and $B = U\mathrm{Diag}(v)U^T$.

Proof. The proof follows similar steps as in [3, Theorem 5]. □

Example 4.15. By way of example, consider the problem of finding the critical angles between SDP cone $\mathbb{S}^n_+$ and the cone

$$\mathcal{D}_n := \{A \in \mathbb{S}^n : \lambda_n(A) \leq \mathrm{tr}(A)\}.$$

n	$\cos\theta$	θ
3	1	0
	$-1/\sqrt{3}$	0.6959π
4	1	0
	$-1/\sqrt{5}$	0.6476π
	$-2/\sqrt{7}$	0.7728π
5	1	0
	$-1/\sqrt{7}$	0.6234π
	$-\sqrt{2}/\sqrt{5}$	0.7180π
	$-3/\sqrt{13}$	0.8128π

Table 4.1: Critical angles between $\mathbb{S}^n_+$ and $\mathcal{D}_n$.

Both cones are orthogonally invariant. The associated permutation invariant cones are $\mathbb{R}^n_+$ and

$$D_n := \{x \in \mathbb{R}^n : \max\{x_1, \ldots, x_n\} \leq x_1 + \ldots + x_n\},$$

respectively. Beware that $\mathbb{S}^n_+$ and $\mathcal{D}_n$ are non-polyhedral cones in $\mathbb{S}^n$, so a direct computation of $\Gamma(\mathbb{S}^n_+, \mathcal{D}_n)$ could be difficult. Computing $\Gamma(\mathbb{R}^n_+, D_n)$ is much easier, because $\mathbb{R}^n_+$ and D_n are simplicial cones in $\mathbb{R}^n$. Table 4.1 is filled with the help of [8, Theorem 7.3].

Remark 4.16. It is possible to derive an explicit formula for the maximal angle between $\mathbb{S}^n_+$ and $\mathcal{D}_n$. One gets

$$\Theta(\mathbb{S}^n_+, \mathcal{D}_n) = \arccos\left(\frac{2-n}{\sqrt{n^2-3n+3}}\right) \tag{4.32}$$

for all $n \geq 2$. For obtaining (4.32) we exploit the fact that D_n is a polyhedral cone generated by n linearly independent unit vectors, namely, the permutations of the vector

$$w = \frac{1}{\sqrt{n^2-3n+3}}\,(2-n, 1, \ldots, 1)^T.$$

4.4.1 The SDP cone versus the cone of nonnegative matrices

We now address the difficult problem of estimating the maximal angle between the SDP cone $\mathbb{S}^n_+$ and the cone

$$\mathcal{N}_n := \{B \in \mathbb{S}^n : B \text{ is nonnegative entrywise}\}.$$

Such problem was raised in a recent paper by Goldberg and Shaked-Monderer [2]. The following facts are known, see [2] for the asymptotic formula stated in Proposition 4.17(c).

Proposition 4.17. *One has:*

(a) $\Theta(\mathbb{S}^2_+, \mathcal{N}_2) = (3/4)\pi$. *Furthermore, the pair of matrices achieving this maximal angle is unique and given by*

$$(A, B) = \left(\begin{bmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{bmatrix}, \begin{bmatrix} 0 & 1/\sqrt{2} \\ 1/\sqrt{2} & 0 \end{bmatrix}\right).$$

(b) $\Theta(\mathbb{S}^n_+, \mathcal{N}_n)$ *is nondecreasing as function of* n*. More generally,*

$$\Gamma(\mathbb{S}^n_+, \mathcal{N}_n) \subseteq \Gamma(\mathbb{S}^{n+1}_+, \mathcal{N}_{n+1}).$$

(c) $\lim_{n\to\infty} \Theta(\mathbb{S}^n_+, \mathcal{N}_n) = \pi$.

The next theorem lists various conditions that are necessary for antipodality in $(\mathbb{S}^n_+, \mathcal{N}_n)$. We start by writing a linear algebra result concerning the smallest eigenvalue of a nonnegative symmetric matrix.

Lemma 4.18. *Let* $B \in \mathcal{N}_n$*. Then* $\sqrt{2}\,\lambda_1(B) + \|B\| \geq 0$*, with equality if and only if*

$$\begin{cases} \lambda_1(B) + \lambda_n(B) = 0, \\ \lambda_2(B) = 0, \ldots, \lambda_{n-1}(B) = 0. \end{cases} \tag{4.33}$$

Proof. In order to alleviate the notation, we write $\lambda_i := \lambda_i(B)$ for all $i \in \{1, \ldots, n\}$. Since B is nonnegative entrywise, the spectral radius

$$\rho(B) := \max_{1\leq i\leq n} |\lambda_i|$$

of B is an eigenvalue of B. Hence, $\rho(B) = \lambda_n \geq -\lambda_1$. On the other hand,

$$[\lambda_1^2 + \ldots + \lambda_n^2]^{1/2} \geq [\lambda_1^2 + \lambda_n^2]^{1/2} \geq \frac{|\lambda_1| + |\lambda_n|}{\sqrt{2}} \geq \frac{\lambda_n - \lambda_1}{\sqrt{2}}.$$

It follows that

$$\|B\| + \sqrt{2}\lambda_1 \geq \left(\frac{\lambda_n + \lambda_1}{\sqrt{2}}\right) \geq 0.$$

This completes the proof of the lemma. □

Theorem 4.19. *Let* $n \geq 3$*. The following conditions are necessary for* (A, B) *to be an antipodal pair of* $(\mathbb{S}^n_+, \mathcal{N}_n)$*:*

(a) A *is not in* $\mathcal{N}_n$ *and* B *is not in* $\mathbb{S}^n_+$.

(b) $B = A^{\mathcal{N}_n} := \|\Pi_{\mathcal{N}_n}(-A)\|^{-1}\Pi_{\mathcal{N}_n}(-A)$.

(c) $A = B^{\mathbb{S}^n_+} := \|\Pi_{\mathbb{S}^n_+}(-B)\|^{-1}\Pi_{\mathbb{S}^n_+}(-B)$.

(d) $B_{i,i} = 0$ *for all* $i \in \{1, \ldots, n\}$.

(e) $AB = BA$.

(f) $\operatorname{rank}(B) \geq 2$.

(g) $\operatorname{rank}(A) = \operatorname{card}\{i : \lambda_i(B) < 0\} \leq n - 2$.

Proof. Part (a). This is because the angle between A and B is at least $(3/4)\pi$.
Part (b). Note that B solves the minimization problem

$$f(A) := \min\{\langle A, Y\rangle : Y \in \mathcal{N}_n, \|Y\| = 1\}.$$

This problem has clearly a unique solution, namely, the matrix $Y = A^{\mathcal{N}_n}$ whose entries are

$$Y_{i,j} = -(1/c)\min\{0, A_{i,j}\}$$

with

$$c := \|\Pi_{\mathcal{N}_n}(-A)\| = \left[\sum_{i,j=1}^{n} (\min\{0, A_{i,j}\})^2\right]^{1/2}.$$

Part (c). Similarly, A solves the minimization problem

$$g(B) := \min\{\langle X, B\rangle : X \in \mathbb{S}^n_+, \|X\| = 1\}. \tag{4.34}$$

We claim that $B^{\mathbb{S}^n_+}$ is the unique solution to (4.34). Since A and B commute, there exist an orthonormal basis $\{x_1, \dots, x_n\}$ of $\mathbb{R}^n$ and a unit vector $\gamma \in \mathbb{R}^n_+$ such that

$$A = \sum_{i=1}^{n} \gamma_i x_i x_i^T, \quad B = \sum_{i=1}^{n} \lambda_i(B) x_i x_i^T. \tag{4.35}$$

One has

$$\begin{aligned} \langle \gamma, \lambda(B)\rangle = \langle A, B\rangle = g(B) &= \min_{\substack{\|\xi\|=1 \\ \xi \geq 0}} \sum_{i=1}^{n} \left\langle \xi_i x_i x_i^T, B\right\rangle \\ &= \min_{\substack{\|\xi\|=1 \\ \xi \geq 0}} \langle \xi, \lambda(B)\rangle. \end{aligned} \tag{4.36}$$

Hence, γ solves the minimization problem (4.36). But such problem admits a unique solution, which can be computed explicitly in terms of the $\lambda_i(B)$'s. One gets

$$\gamma_i = -(1/d)\min\{0, \lambda_i(B)\}, \tag{4.37}$$

with

$$d := \left[\sum_{i=1}^{n} (\min\{0, \lambda_i(B)\})^2\right]^{1/2}.$$

By combining (4.35) and (4.37) one sees that, up to normalization, A is the projection of $-B$ onto $\mathbb{S}^n_+$.
Part (d). This is a consequence of (b).
Part (e). This is a consequence of Proposition 4.13.
Part (f). As a consequence of (d), at least two eigenvalues of B are different from 0.
Part (g). Let r be the number of negative eigenvalues of B. Then, thanks to (4.35) and (4.37), one has

$$A = -(1/d)\sum_{i=1}^{r} \lambda_i(B) x_i x_i^T$$

with $d = [\sum_{k=1}^{r} \lambda_k^2(B)]^{1/2}$. In particular, the rank$(A) = r$. In the remaining part of the proof, we use the notation $\lambda_i := \lambda_i(B)$. One has rank$(A) \leq n-1$, because A must be on the boundary of $\mathbb{S}_+^n$. Suppose that rank$(A) = n-1$. We must arrive to a contradiction. From the proof of (c), one sees that $\lambda_i < 0$ for all $i \in \{1, \ldots, n-1\}$ and

$$\langle A, B\rangle = -\left(\lambda_1^2 + \ldots + \lambda_{n-1}^2\right)^{1/2}.$$

On the other hand, one has

$$\lambda_1 + \ldots + \lambda_n = 0, \quad \lambda_1^2 + \ldots + \lambda_n^2 = 1, \quad \lambda_n > 0.$$

One gets in this way

$$\lambda_n = \frac{1}{\sqrt{2}}\left[1 + 2\sum_{1\leq i<j\leq n-1} \lambda_i\lambda_j\right]^{1/2} > \frac{1}{\sqrt{2}} \tag{4.38}$$

and $\langle A, B\rangle = -[1-\lambda_n^2]^{1/2} > -1/\sqrt{2}$, contradicting the inequality $\Theta(\mathbb{S}_+^n, \mathcal{N}_n) \geq (3/4)\pi$. □

The next corollary fully settles the case $n = 3$.

Corollary 4.20. *One has* $\Theta(\mathbb{S}_+^3, \mathcal{N}_3) = (3/4)\pi$. *Furthermore,* (A, B) *is an antipodal pair of* $(\mathbb{S}_+^3, \mathcal{N}_3)$ *if and only if*

$$A = xx^T, \quad B = \frac{1}{\sqrt{2}}(yy^T - xx^T)$$

with $x, y \in \mathbb{R}^3$ *such that*

$$\begin{cases} \|x\| = 1, \ \|y\| = 1, \ \langle x, y\rangle = 0, \\ y_iy_j \geq x_ix_j \ \ \textit{for} \ \ 1 \leq i \leq j \leq 3. \end{cases} \tag{4.39}$$

Proof. Let (A, B) be an antipodal pair of $(\mathbb{S}_+^3, \mathcal{N}_3)$. Theorem 4.19(g) implies that rank$(A) = 1$. Hence, $A = xx^T$ with $\|x\| = 1$. By using Lemma 4.18 one gets

$$\begin{aligned} \cos[\Theta(\mathbb{S}_+^3, \mathcal{N}_3)] &= \min_{\|u\|=1} \min_{\substack{B\in\mathcal{N}_3 \\ \|B\|=1}} \langle uu^T, B\rangle \\ &= \min_{\substack{B\in\mathcal{N}_3 \\ \|B\|=1}} \lambda_1(B) = -1/\sqrt{2}. \end{aligned}$$

The second part of the corollary is obtained by using (4.33). □

Remark 4.21. If t, s are nonnegative reals such that $t^2 + s^2 = 1$, then

$$x = (1/\sqrt{2})\,(t, s, -1)^T, \quad y = (1/\sqrt{2})\,(t, s, 1)^T$$

satisfy (4.39) and lead to the antipodal pair

$$(A, B) = \left(\frac{1}{2}\begin{bmatrix} t^2 & ts & -t \\ ts & s^2 & -s \\ -t & -s & 1 \end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 0 & t \\ 0 & 0 & s \\ t & s & 0 \end{bmatrix}\right).$$

Hence, the number of antipodal pairs of $(\mathbb{S}_+^3, \mathcal{N}_3)$ is not finite, not even countable.

From the proof of Theorem 4.19 one sees that

$$\cos\left[\Theta(\mathbb{S}_+^n, \mathcal{N}_n)\right] = \min\{f(A) : A \in \mathbb{S}_+^n, \|A\| = 1\} \quad (4.40)$$
$$= \min\{g(B) : B \in \mathcal{N}_n, \|B\| = 1\} \quad (4.41)$$

with

$$f(A) = -\left[\sum_{i,j=1}^{n} \left(\min\{0, A_{i,j}\}\right)^2\right]^{1/2},$$
$$g(B) = -\left[\sum_{i=1}^{n} \left(\min\{0, \lambda_i(B)\}\right)^2\right]^{1/2}.$$

The minimization problems (4.40) and (4.41) are hard to solve in practice, because they are nonconvex and nonsmooth. However, the variational formulas (4.40) and (4.41) are useful to obtain lower bounds for $\Theta(\mathbb{S}_+^n, \mathcal{N}_n)$.

Example 4.22. Consider for instance the case $n = 5$. The nonnegative matrix B

$$B = \frac{1}{\sqrt{10}} \begin{bmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \end{bmatrix}$$

has unit norm and its eigenvalues are

$$\lambda_1(B) = \lambda_2(B) = \frac{-1-\sqrt{5}}{2\sqrt{10}}, \quad \lambda_3(B) = \lambda_4(B) = \frac{-1+\sqrt{5}}{2\sqrt{10}}, \quad \lambda_5(B) = \frac{2}{\sqrt{10}}.$$

Hence

$$g(B) = -\left[\left(\frac{-1-\sqrt{5}}{2\sqrt{10}}\right)^2 + \left(\frac{-1-\sqrt{5}}{2\sqrt{10}}\right)^2\right]^{1/2} = -\frac{1+\sqrt{5}}{2\sqrt{5}}$$

and

$$\Theta(\mathbb{S}_+^5, \mathcal{N}_5) \geq \arccos\left(-\frac{1+\sqrt{5}}{2\sqrt{5}}\right) \approx 0.7575\pi.$$

The next proposition is a complement to Theorem 4.19. It applies to the case $n \geq 5$ only.

Proposition 4.23. *Suppose that $n \geq 5$. Let (A, B) be an antipodal pair of $(\mathbb{S}_+^n, \mathcal{N}_n)$. Then,*

$$\lambda_2(B) < 0 < \lambda_{n-1}(B).$$

In particular, $\operatorname{rank}(B) \geq 4$ *and* $\operatorname{rank}(A) \geq 2$.

Proof. Let $\lambda_i := \lambda_i(B)$ for all $i \in \{1, \ldots, n\}$. Suppose that $\lambda_{n-1} \leq 0$. We must arrive to a contradiction. One has $\lambda_i \leq 0$ for all $i \in \{1, \ldots, n-1\}$. The inequality in (4.38) is not strict, but holds in the form " $\geq$". One gets in such a case

$$\langle A, B \rangle = -\sqrt{1 - \lambda_n^2} \ \geq -1/\sqrt{2}\,,$$

which contradicts the inequality $\Theta(\mathbb{S}_+^n, \mathcal{N}_n) > (3/4)\pi$, cf. Example 4.22. Hence, $\lambda_{n-1} > 0$. If $\lambda_2 \geq 0$, then

$$-\lambda_1 = \lambda_2 + \ldots + \lambda_n \geq \lambda_{n-1} + \lambda_n > \lambda_n,$$

contradicting the fact that $\lambda_n = \rho(B)$. □

It is quite difficult to obtain an explicit formula for $\Theta(\mathbb{S}_+^n, \mathcal{N}_n)$ when $n \geq 4$. Intensive numerical testing suggests that $\Theta(\mathbb{S}_+^4, \mathcal{N}_4)$ is equal to $(3/4)\pi$, but we do not have a formal proof of this fact. From Example 4.22 one knows that

$$\Theta(\mathbb{S}_+^n, \mathcal{N}_n) > (3/4)\pi \quad \text{for all } n \geq 5.$$

Some words on the numerical estimation of $\Theta(\mathbb{S}_+^n, \mathcal{N}_n)$ are in order. As a consequence of [8, Theorem 3.2], one can write

$$\Theta(\mathbb{S}_+^n, \mathcal{N}_n) \ = \ 2 \arccos \sqrt{2t_n}\,,$$

where t_n denotes the optimal value of the nonlinear program

$$\begin{cases} \text{minimize } f_0(Z,t) := t, \\ (Z,t) \in \mathbb{S}^n \times \mathbb{R}, \\ f_1(Z,t) := \frac{1}{2}\left[\operatorname{dist}(Z, \mathbb{S}_+^n)\right]^2 - t \leq 0, \\ f_2(Z,t) := \frac{1}{2}\left[\operatorname{dist}(Z, -\mathcal{N}_n)\right]^2 - t \leq 0, \\ f_3(Z,t) := \|Z\|^2 - 1 = 0. \end{cases} \tag{4.42}$$

The gradients of f_0, f_1, f_2, and f_3, are all easily computable. For instance, the partial gradients of f_1 and f_2 with respect to Z are given by

$$\begin{aligned} \langle \nabla_Z f_1(Z,t), D \rangle &= D - \Pi_{\mathbb{S}_+^n}(D), \\ \langle \nabla_Z f_2(Z,t), D \rangle &= D + \Pi_{\mathcal{N}_n}(-D). \end{aligned}$$

Projecting onto $\mathbb{S}_+^n$ and $\mathcal{N}_n$ offers no difficulty. Table 4.2 has been filled by solving (4.42) with the help of the package "fmincon" of MATLAB. This is done for each $n \in \{4, \ldots, 30\}$. Since (4.42) is a nonconvex optimization problem, we are not sure if "fmincon" is yielding a global solution or just a local solution. For this reason we are rather conservative and consider the figures in Table 4.2 only as lower bounds for $\Theta(\mathbb{S}_+^n, \mathcal{N}_n)$. These figures have been rounded down to four decimal places.

n	$\Theta(\mathbb{S}_+^n, \mathcal{N}_n)$	n	$\Theta(\mathbb{S}_+^n, \mathcal{N}_n)$	n	$\Theta(\mathbb{S}_+^n, \mathcal{N}_n)$
4	$0.7500\,\pi$	13	$0.7649\,\pi$	22	$0.7719\,\pi$
5	$0.7575\,\pi$	14	$0.7658\,\pi$	23	$0.7722\,\pi$
6	$0.7575\,\pi$	15	$0.7677\,\pi$	24	$0.7735\,\pi$
7	$0.7575\,\pi$	16	$0.7699\,\pi$	25	$0.7735\,\pi$
8	$0.7607\,\pi$	17	$0.7699\,\pi$	26	$0.7735\,\pi$
9	$0.7607\,\pi$	18	$0.7699\,\pi$	27	$0.7739\,\pi$
10	$0.7609\,\pi$	19	$0.7703\,\pi$	28	$0.7750\,\pi$
11	$0.7626\,\pi$	20	$0.7719\,\pi$	29	$0.7750\,\pi$
12	$0.7649\,\pi$	21	$0.7719\,\pi$	30	$0.7753\,\pi$

Table 4.2: Lower bound for $\Theta(\mathbb{S}_+^n, \mathcal{N}_n)$.

For the sake of completeness, we record below a recent result due to Goldberg and Shaked-Monderer [2].

Proposition 4.24. *Consider a dimension n of the form $n = (q+1)(q^3+1)$, with q being a prime power. Then*

$$\Theta(\mathbb{S}_+^n, \mathcal{N}_n) \geq \arccos\left(-\frac{\sqrt{q^2+1}}{q+1}\right). \tag{4.43}$$

The lower bound (4.43) has the merit of being easily computable, but it applies only to special choices of n and it is not optimal in general. Consider for instance the choice $q = 2$, which corresponds to the first prime power. The inequality (4.43) becomes

$$\Theta(\mathbb{S}_+^{27}, \mathcal{N}_{27}) \geq \arccos\left(-\sqrt{5}/3\right) \approx 0.7677\,\pi,$$

but Table 4.2 yields the better lower bound $\Theta(\mathbb{S}_+^{27}, \mathcal{N}_{27}) \geq 0.7739\,\pi$.

Bibliography

[1] M. Fiedler and E. Haynsworth. Cones which are topheavy with respect to a norm. *Linear and Multilinear Algebra*, 1 (1973), 203–211.

[2] F. Goldberg and N. Shaked-Monderer. On the maximal angle between copositive matrices. July 2013, submitted. Temporarily available at http://arxiv.org/pdf/1307.7519.pdf.

[3] A. Iusem and A. Seeger. Angular analysis of two classes of non-polyhedral convex cones: the point of view of optimization theory. *Comput. Applied Math.*, 26 (2007), 191–214.

[4] Y. Lyubich. Perron-Frobenius theory for finite-dimensional spaces with a hyperbolic cone. *Linear Algebra Appl.* 220 (1995), 283–309.

[5] D.G. Obert. The angle between two cones. *Linear Algebra Appl.*, 144 (1991), 63–70.

[6] A. Seeger. Eigenvalue analysis of equilibrium processes defined by linear complementarity conditions. *Linear Algebra Appl.*, 292 (1999), 1–14.

[7] A. Seeger. Epigraphical cones I. *J. Convex Analysis*, 18 (2011), 1171–1196.

[8] A. Seeger and D. Sossa. Critical angles bewteen two convex cones. Part I: General theory. Submitted to this journal.

[9] M. Tenenhaus. Canonical analysis of two convex polyhedral cones and applications. *Psychometrika* 53 (1988), 503–524.

Chapter 5

On the central path in symmetric cone linear programming

Héctor Ramírez[1] - David Sossa[2]

Abstract. In this paper we study the convergence and the limiting behavior of the central path in symmetric cone linear programming.

Mathematics Subject Classification. 17C27, 90C05
Key words. Euclidean Jordan algebra, symmetric cone linear programming, central path

5.1 Introduction

The concept of the central path plays a fundamental role in the study of interior point algorithms. This subject has been widely studied in the context of Semidefinite Programming (SDP). For instance, it is known that the central path always converge and that the limit point lies on the relative interior of the optimal set (cf. [9, 6]). Furthermore, it was shown in [9] that under the strictly complementarity assumption the central path converges to the analytic center of the optimal set. In the general case the result it is not true, but the limit point can be still characterized by mean of an optimization problem (cf. [7]).

In this paper we consider the central path in a more general problem known as Symmetric Cone Linear Programming (SCLP). The SCLP not only includes the SDP, but also Linear

[1]Universidad de Chile, Center of Mathematical Modeling, Casilla 170-3, Santiago, Chile. E-mail: hramirez@dim.uchile.cl.
[2]Universidad de Chile, Department of Mathematical Engineering, Casilla 170-3, Santiago, Chile. E-mail: dsossa@dim.uchile.cl.

and Second Order Cone Programming. As in SDP, the importance of the central path in SCLP lies in the study of interior point algorithms for solving the SCLP. We refer the papers [11, 12] of Nesterov and Todd where the interior point algorithms were designed for the first time for solving SCLP. In [3, 4], Faybusovich shows that Jordan algebra techniques can be used to describe Nesterov-Todd's method and some others interior point algorithms in a clear way. After this, important efforts were focused on the development of new algorithms for solving SCLP (e.g. [13, 14, 15]).

The aim of this paper is to extend some basic results of the central path of SDP to the central path of SCLP. Our analysis focuses on the convergence and the limiting behavior of the central path of SCLP under some standard assumptions.

The outline of this paper is as follows. In Section 2 we provide some preliminary materials concerning to Euclidean Jordan algebras and we introduce the central path in SCLP. In Section 3 we study the convergence of the central path. Finally, the limiting behavior is analyzed in Section 4.

5.2 Preliminaries

5.2.1 Euclidean Jordan algebras

Throughout this paper one assumes that $(\mathbb{V}, \circ, \langle\cdot,\cdot\rangle)$ is a Euclidean Jordan algebra (EJA) with unit element $e \in \mathbb{V}$. This means that $\mathbb{V}$ is a finite dimensional real vector space equipped with an inner product $\langle\cdot,\cdot\rangle$ and a bilinear function $\circ : \mathbb{V} \times \mathbb{V} \to \mathbb{V}$ satisfying the axioms:

$$\begin{cases} x \circ y = y \circ x & \text{for all } x, y \in \mathbb{V}, \\ x \circ (x^2 \circ y) = x^2 \circ (x \circ y) & \text{for all } x, y \in \mathbb{V}, \\ \langle x \circ y, z\rangle = \langle y, x \circ z\rangle & \text{for all } x, y, z \in \mathbb{V}, \\ e \circ x = x & \text{for all } x \in \mathbb{V}. \end{cases}$$

Here and in the sequel one uses the notation $x^2 = x \circ x$. Higher order powers are defined recursively by $x^{k+1} = x \circ x^k$. We list below some definitions and properties concerning to the general theory of Euclidean Jordan algebras that we use throughout this work.

- The *rank* of $\mathbb{V}$ is declared to be

$$r = \max\{\deg(x) : x \in \mathbb{V}\},$$

  where $\deg(x)$ is the smallest positive integer k such that $\{e, x, x^2, \ldots, x^k\}$ is linearly dependent.

- An *ideal* of the EJA algebra $\mathbb{V}$ is a subalgebra $I \subseteq \mathbb{V}$ such that $x \circ u \in I$ whenever $x \in \mathbb{V}$ and $u \in I$. An EJA is *simple* if it does not contain any nontrivial ideal.

- The set of square elements of $\mathbb{V}$ defined as

$$\mathcal{K} := \{x^2 : x \in \mathbb{V}\},$$

is a *symmetric cone* (cf. [2, Theorem III.2.1]). This mean that $\mathcal{K}$ is a self-dual closed convex cone with nonempty interior and for any two elements $x, y \in \text{int}\mathcal{K}$, there exists an invertible linear transformation $\Gamma : \mathbb{V} \to \mathbb{V}$ such that $\Gamma(\mathcal{K}) = \mathcal{K}$ and $\Gamma(x) = y$.

- An element $c \in \mathbb{V}$ is an *idempotent* if $c^2 = c$.

- An idempotent c is *primitive* if it is nonzero and cannot be written as a sum of two nonzero idempotents.

- A *Jordan frame* is a collection $\{c_1, \ldots, c_r\}$ of primitive idempotents satisfying

$$\sum_{i=1}^{r} c_i = e \quad \text{and} \quad c_i \circ c_j = 0 \ \text{ when } i \neq j.$$

We recall below a spectral decomposition theorem taken from [2, Theorem III.1.2].

Theorem 5.1. *Let $(\mathbb{V}, \circ, \langle\cdot,\cdot\rangle)$ be an EJA with rank r. Then, for every $x \in \mathbb{V}$, there exists a Jordan frame $\{c_1, \ldots, c_r\}$ and real numbers $\lambda_1, \ldots, \lambda_r$ such that*

$$x = \lambda_1 c_1 + \ldots + \lambda_r c_r.$$

The λ_i's are uniquely determined by x and they are called the eigenvalues *of x.*

- Let $x \in \mathbb{V}$ with eigenvalues $\lambda_1, \ldots, \lambda_r$. We define:

$$\begin{aligned} &\textit{trace of } x \ : \ \text{tr}(x) = \lambda_1 + \ldots + \lambda_r, \\ &\textit{determinant of } x \ : \ \det(x) = \Pi_{i=1}^{r} \lambda_i, \\ &\textit{rank of } x \ : \ \text{rank}(x) = \text{number of nonzero eigenvalues of } x. \end{aligned}$$

- For $x \in \mathbb{V}$ the *Lyapunov operator* associated with a given x is the linear map $L_x : \mathbb{V} \to \mathbb{V}$ given by $L_x y = x \circ y$ and the *quadratic representation* of x is the linear map $P : \mathbb{V} \to \mathbb{V}$ given by $P_x = 2L_x^2 - L_{x^2}$.

- One says that two elements $a, b \in \mathbb{V}$ *operator commute* if

$$a \circ (b \circ z) = b \circ (a \circ z) \quad \text{for all } z \in \mathbb{V}.$$

From [13, Theorem 27], this is equivalent to the existence of a Jordan frame $\{c_1, \ldots, c_r\}$ and real numbers $\lambda_1, \ldots, \lambda_r$ and $\mu_1, \ldots, \mu_r$ such that

$$\begin{aligned} a &= \lambda_1 c_1 + \ldots + \lambda_r c_r \\ b &= \mu_1 c_1 + \ldots + \mu_r c_r. \end{aligned}$$

- An element $x \in \mathbb{V}$ is said to be *invertible* if there exists an element $y \in \mathbb{V}$ such that it is a linear combination of powers of x and satisfies $x \circ y = e$. The element y is unique. It is called the *inverse* of x and is denoted by $y = x^{-1}$. Note that if an invertible element x admits the spectral decomposition $x = \sum_{i=1}^{r} \lambda_i c_i$ then its inverse admits the spectral decomposition $x^{-1} = \sum_{i=1}^{r} (1/\lambda_i) c_i$. In particular, x and x^{-1} operator commute.

- Let $c \in \mathbb{V}$ be an idempotent element. The EJA $\mathbb{V}$ can be decomposed into three mutually orthogonal subspaces:

$$\mathbb{V} = \mathbb{V}(c, 1) \oplus \mathbb{V}(c, 1/2) \oplus \mathbb{V}(c, 0),$$

  where

$$\mathbb{V}(c, \ell) := \{x \in \mathbb{V} : c \circ x = \ell x\}$$

  with $\ell = 1, 1/2, 0$. This decomposition is called the *Peirce decomposition* of $\mathbb{V}$ with respect to c. The projections in the Peirce decomposition are the following:

$$\begin{aligned} &\text{onto } \mathbb{V}(c, 1) \ : \ P_c, \\ &\text{onto } \mathbb{V}(c, 1/2) \ : \ \mathcal{I} - P_c - P_{e-c}, \\ &\text{onto } \mathbb{V}(c, 0) \ : \ P_{e-c}, \end{aligned}$$

  where $\mathcal{I}$ stands for the identity operator in $\mathbb{V}$ (cf. [2, Chapter IV]).

Example 5.2. Typical examples of Euclidean Jordan algebras are:

(a) *Euclidean Jordan algebra of n-dimensional vectors:*

$$\mathbb{V} = \mathbb{R}^n, \quad x \circ y = x \odot y, \quad \langle x, y \rangle = \sum_{i=1}^{n} x_i y_i, \quad r = n, \quad \mathcal{K} = \mathbb{R}^n_+,$$

where "$\odot$" denotes the componentwise (Hadamard) product of vectors x and y. Here, the unitary element is $e = (1, \ldots, 1) \in \mathbb{R}^n$.

(b) *Euclidean Jordan algebra of quadratic forms*:

$$\mathbb{V} = \mathbb{R}^n, \quad x \circ y = (\langle x, y \rangle, x_1 \bar{y} + y_1 \bar{x}), \quad \langle x, y \rangle = \sum_{i=1}^{n} x_i y_i, \quad r = 2, \quad \mathcal{K} = \mathcal{L}_n,$$

where $\bar{x} = (x_2 \ldots, x_n) \in \mathbb{R}^{n-1}$ and $\mathcal{L}_n := \{x \in \mathbb{R}^n : \|\bar{x}\| \leq x_1\}$ is the second-order cone in $\mathbb{R}^n$. Here the unit element is $e = (1, 0 \ldots, 0) \in \mathbb{R}^n$.

(c) *Euclidean Jordan algebra of n-dimensional symmetric matrices:*

$$\mathbb{V} = \mathbb{S}^n, \quad X \circ Y = (XY + YX)/2, \quad \langle X, Y \rangle = \mathrm{Tr}(XY), \quad r = n, \quad \mathcal{K} = \mathbb{S}^n_+.$$

Here $\mathbb{S}^n$ denotes the space of symmetric matrices of order n, $\mathbb{S}^n_+$ stands for the cone of positive semidefinite symmetric matrices in $\mathbb{S}^n$ and "Tr" denotes the trace of a matrix. In this setting, the identity matrix $I \in \mathbb{S}^n$ is the unit element e.

5.2.2 Symmetric cone linear programming

The primal SCLP is formulated as

$$(P) \quad \min_{x\in\mathbb{V}}\{\langle c, x\rangle : \mathcal{A}x = b,\ x \in \mathcal{K}\}$$

where $(\mathbb{V}, \circ, \langle\cdot,\cdot\rangle)$ is a Euclidean Jordan algebra, $\mathcal{K}$ is the cone of squares elements in $\mathbb{V}$, $c \in \mathbb{V}$, $b \in \mathbb{R}^m$ and $\mathcal{A} : \mathbb{V} \to \mathbb{R}^m$ is a linear map. The dual problem associated with (P) is given by

$$(D) \quad \max_{(y,s)\in\mathbb{R}^m\times\mathbb{V}}\{b^T y : \mathcal{A}^* y + s = c,\ s \in \mathcal{K}\},$$

where $\mathcal{A}^*$ is the adjoint map of $\mathcal{A}$ and the superscript "T" stands for the transpose of a vector. Consequently, $b^T y = \sum_{i=1}^m b_i y_i$. The primal and dual feasible sets are denoted by $\mathcal{F}_P$ and $\mathcal{F}_D$, respectively, whereas $\mathcal{S}_P$ and $\mathcal{S}_D$ denote their respective optimal sets.

Remark 5.3. Without loss of generality, we can think that the space $(\mathbb{V}, \langle\cdot,\cdot\rangle)$ on problem (P) is an arbitrary Euclidean space and the cone $\mathcal{K}$ is an arbitrary symmetric cone on $\mathbb{V}$. This follows from [2, Theorem III.3.1] which establishes that it is possible to construct a Jordan product " $\circ$ " such that $(\mathbb{V}, \circ, \langle\cdot,\cdot\rangle)$ is a Euclidean Jordan algebra and $\mathcal{K}$ coincides with the cone of square elements on $\mathbb{V}$.

Example 5.4. Linear programming, second order cone programming and semidefinite programming are particular cases of SCLP. In fact, this is obtained by taking, respectively, $\mathcal{K} = \mathbb{R}^n_+$, $\mathcal{K} = \mathcal{L}_n$ and $\mathcal{K} = \mathbb{S}^n_+$ with their corresponding spaces as in Example 5.2.

Throughout this paper, we take the following standard assumptions:

(A1) $\mathcal{A}$ is a surjective linear map;

(A2) $\mathcal{F}_P \cap \text{int}\mathcal{K} \neq \emptyset$ and $\mathcal{F}_D \cap (\mathbb{R}^m \times \text{int}\mathcal{K}) \neq \emptyset$;

(A3) $(\mathbb{V}, \circ, \langle\cdot,\cdot\rangle)$ is simple.

Assumption (A1) ensures that the dual variables y and s are in one-to-one correspondence. Assumption (A2) implies that $\mathcal{S}_P$ and $\mathcal{S}_D$ are nonempty and bounded and there is no duality gap between (P) and (D). Assumption (A3) ensures that the set of primitive idempotents elements of $\mathbb{V}$ is compact (cf. [8]) and that the EJA is *scalarizable* (cf. [2, Proposition III.4.1]), i.e., there exists a positive constant ρ such that

$$\langle u, v\rangle = \rho\text{tr}(u \circ v) \quad \text{for all } u, v \in \mathbb{V}.$$

The KKT optimality conditions state that $(x, y, s) \in \mathcal{S}_P \times \mathcal{S}_D$ if and only if it solves the following system of equations

$$\mathcal{A}x = b, \quad x \in \mathcal{K}, \tag{5.1}$$
$$\mathcal{A}^* y + s = c, \quad s \in \mathcal{K}, \tag{5.2}$$
$$x \circ s = 0. \tag{5.3}$$

We study the solution of problem (P) by using a penalization scheme with respect to the logarithmic barrier function.

$$(P_\mu) \quad \min_{x\in\mathbb{V}}\{\langle c,x\rangle - \mu \ln\det x : \mathcal{A}x = b,\ x\in \mathrm{int}\mathcal{K}\},\ \ \mu>0.$$

One can show that the dual problem associated with (P_μ) is given by

$$(D_\mu) \quad \max_{(y,s)\in\mathbb{R}^m\times\mathbb{V}}\{\langle b,y\rangle + \mu \ln\det s : \mathcal{A}^*y + s = c,\ s\in \mathrm{int}\mathcal{K}\}$$

The assumption (A2) and the fact that the function $-\ln\det(\cdot)$ is strictly convex over $\mathrm{int}\mathcal{K}$ (cf. [11]) imply that for every $\mu>0$ problem (P_μ) admits a unique solution $x(\mu)$ and the problem (D_μ) admits a unique solution $(y(\mu),s(\mu))$ (cf. [4]). By using the KKT optimality conditions one gets that for every $\mu>0$, the triplet $(x(\mu),y(\mu),s(\mu))$ can be characterized as the unique solution of the system

$$\begin{aligned} \mathcal{A}x &= b, \quad x\in\mathrm{int}\mathcal{K},\\ \mathcal{A}^*y + s &= c, \quad s\in\mathrm{int}\mathcal{K},\\ s &= (\mu/\rho)x^{-1}. \end{aligned} \tag{5.4}$$

It is well-known (cf. [3]) that this set of condition is equivalent to

$$\mathcal{A}x = b, \quad x\in\mathrm{int}\mathcal{K}, \tag{5.5}$$

$$\mathcal{A}^*y + s = c, \quad s\in\mathrm{int}\mathcal{K}, \tag{5.6}$$

$$x\circ s = (\mu/\rho)e, \tag{5.7}$$

The relation (5.7) is known as the *centering condition* of the central path. We refer to the trajectory $\{(x(\mu),y(\mu),s(\mu)) : \mu>0\}$ as the primal-dual central path (or simply central path) in SCLP.

5.3 Convergence of the central path

We start by recalling a trivial but useful lemma that is known as the *orthogonality relation* and it is obtained from the feasibility conditions of (P) and (D).

Lemma 5.5. *Let $(x,y,s),(\tilde{x},\tilde{y},\tilde{s})\in\mathcal{F}_P\times\mathcal{F}_D$. Then the following orthogonality relation is satisfied:*

$$\langle x-\tilde{x}, s-\tilde{s}\rangle = 0. \tag{5.8}$$

Proof. Let $(x,y,s),(\tilde{x},\tilde{y},\tilde{s})\in\mathcal{F}_P\times\mathcal{F}_D$. That is, they satisfy

$$\begin{cases} \mathcal{A}x = b \\ \mathcal{A}^*y + s = c \end{cases}, \quad \begin{cases} \mathcal{A}\tilde{x} = b \\ \mathcal{A}^*\tilde{y} + \tilde{s} = c. \end{cases}$$

Hence,

$$\langle x-\tilde{x}, s-\tilde{s}\rangle = \langle x-\tilde{x}, \mathcal{A}^*(\tilde{y}-y)\rangle = \langle \mathcal{A}x - \mathcal{A}\tilde{x}, \tilde{y}-y\rangle = 0$$

□

The next lemma shows the boundeness of the central path. It is a generalization of [9, Lemma 3.1].

Lemma 5.6. *Given $\bar{\mu} > 0$, the set $\{(x(\mu), y(\mu), s(\mu)) : 0 < \mu \leq \bar{\mu}\}$ is bounded and all its limit points are in $\mathcal{S}_P \times \mathcal{S}_D$.*

Proof. Let $\bar{\mu} > 0$. Note that $y(\mu) = (\mathcal{A}\mathcal{A}^*)^{-1}\mathcal{A}(c - s(\mu))$. Therefore, it is enough to prove that $\{(x(\mu), s(\mu)) : 0 < \mu \leq \bar{\mu}\}$ is bounded. Let $x^0 \in \mathcal{F}_P$ and $(y^0, s^0) \in \mathcal{F}_D$ be such that $x^0, s^0 \in \text{int}\mathcal{K}$. The orthogonality relation (5.8) implies that

$$\langle x(\mu), s^0\rangle + \langle x^0, s(\mu)\rangle = \langle x(\mu), s(\mu)\rangle + \langle x^0, s^0\rangle. \tag{5.9}$$

From the centering condition (5.7) one has that $\langle x(\mu), s(\mu)\rangle = \rho r \mu$. Hence, for every $\mu \leq \bar{\mu}$ from (5.9) one gets

$$\langle x(\mu), s^0\rangle + \langle x^0, s(\mu)\rangle \leq \gamma := \rho r \bar{\mu} + \langle x^0, s^0\rangle.$$

This implies that

$$\{(x(\mu), s(\mu)) : 0 < \mu \leq \bar{\mu}\} \subseteq \{(x, s) \in \mathcal{K} \times \mathcal{K} : \langle (x, s), \gamma^{-1}(x^0, s^0)\rangle \leq 1\}.$$

Then from [2, Corollary I.1.6] one concludes that $\{(x(\mu), y(\mu), s(\mu)) : 0 < \mu \leq \bar{\mu}\}$ is bounded.

Next, we prove that all the limit points are in $\mathcal{S}_P \times \mathcal{S}_D$. Let (x^*, y^*, s^*) be a limit point of the central path. It is possible to construct a sequence (μ_k) such that

$$(\mu_k, x(\mu_k), y(\mu_k), s(\mu_k)) \to (0, x^*, y^*, s^*),$$

when $k \to +\infty$. We define the map $\Phi : \mathbb{R} \times \mathbb{V} \times \mathbb{R}^m \times \mathbb{V} \to \mathbb{R}^m \times \mathbb{V} \times \mathbb{V}$ as

$$\Phi(\mu, x, y, s) = (\mathcal{A}x - b, \mathcal{A}^*y + s - c, x \circ s - (\mu/\rho)e).$$

Observe that $\Phi(\mu_k, x(\mu_k), y(\mu_k), s(\mu_k)) = (0, 0, 0)$ for every k. Hence, by continuity of Φ we conclude that $\Phi(0, x^*, y^*, s^*) = (0, 0, 0)$. It means $(x^*, y^*, s^*) \in \mathcal{S}_P \times \mathcal{S}_D$. □

In the following proposition we show the convergence of the central path. The arguments are similar to those used by Halická et al. [6] for the case of SDP.

Definition 5.7. *A subset $\mathcal{V} \subset \mathbb{R}^\ell$ is called an* algebraic set *if $\mathcal{V}$ is described as*

$$\mathcal{V} = \{z \in \mathbb{R}^\ell : p_1(z) = 0, \ldots, p_r(z) = 0\},$$

where p_i are polynomial functions on $\mathbb{R}^\ell$.

We known that a matrix in $\mathbb{S}^n$ is positive definite if and only if their leading principal minors are positive. We explain how this propertie is extended to the general context of EJA (cf. [2, Section VI.3]). Let $\{c_1, \ldots, c_r\}$ be any Jordan frame. We set $\mathbb{V}^{(k)} := \mathbb{V}(c_1 + \ldots + c_k, 1)$, $k \in \{1, \ldots, r\}$ which are subalgebras. One can see that

$$\mathbb{V}^{(1)} \subset \mathbb{V}^{(2)} \subset \ldots \subset \mathbb{V}^{(r)} = \mathbb{V}.$$

Let P_k be the orthogonal projection onto $\mathbb{V}^{(k)}$. The principal minor Δ_k is the polynomial function defined on $\mathbb{V}$ by

$$\Delta_k(x) = \det{}^{(k)}(P_k(x)),$$

where $\det^{(k)}$ denotes the determinant with respect to the subalgebra $\mathbb{V}^{(k)}$. It is known that

$$x \in \text{int}\mathcal{K} \Leftrightarrow \Delta_k(x) > 0, \text{ for all } k = 1, \ldots, r,$$

Proposition 5.8. *The central path in symmetric cone linear programming always converges.*

Proof. Let (x^*, y^*, s^*) be the limit point of the primal-dual central path. We define the sets

$$\mathcal{V} := \left\{ (\bar{x}, \bar{y}, \bar{s}, \mu) \;\middle|\; \begin{array}{r} \mathcal{A}\bar{x} = 0 \\ \mathcal{A}^*\bar{y} + \bar{s} = 0 \\ (\bar{x} + x^*) \circ (\bar{s} + s^*) - \mu e = 0 \end{array} \right\}$$

$$\mathcal{U} := \{(\bar{x}, \bar{y}, \bar{s}, \mu) : \Delta_k(\bar{x} + x^*) > 0,\ \Delta_k(\bar{s} + s^*) > 0,\ \mu > 0, k = 1, \ldots, r\}.$$

Let $n := \dim\mathbb{V}$. Note that, for a fixed basis of $\mathbb{V}$, the set $\mathcal{V}$ can be seen as the locus of common zeros of a collection of polynomial functions on $\mathbb{R}^{n+m+n+1}$. That is $\mathcal{V}$ is an algebraic set. Analogously, the set $\mathcal{U}$ can be seen as defined by inequalities of polynomial functions on $\mathbb{R}^{n+m+n+1}$. The set $\mathcal{V} \cap \mathcal{U}$ corresponds to the point $(\bar{x}, \bar{y}, \bar{s}, \mu)$ such that $x(\mu) = \bar{x} + x^*$ and $s(\mu) = \bar{s} + s^*$, for all $\mu > 0$. Moreover the zero element is in the closure of $\mathcal{V} \cap \mathcal{U}$, by construction. Hence, we can use the curve selection lemma [6, Lemma A.2] in order to conclude the result as in [6, Theorem A.3]. The details are omitted. □

5.4 Limiting behavior of the central path

In this section we study some properties of the limit point of the central path in SCLP. This is done by using the Peirce decomposition of $\mathbb{V}$ with respect to a particular idempotent element.

In what follows, we suppose that (x^*, y^*, s^*) is the limit point of the central path in SCLP. Note that the relation (5.4) implies that for every $\mu > 0$, $x(\mu)$ and $s(\mu)$ operator commute. Hence, they admit a simultaneous spectral decomposition, namely,

$$x(\mu) = \sum_{i=1}^{r} \gamma_i(\mu) c_i(\mu), \qquad s(\mu) = \sum_{i=1}^{r} \delta_i(\mu) c_i(\mu). \tag{5.10}$$

On the other hand, from the optimality conditions (5.1)-(5.3) one gets

$$x^* \in \mathcal{K},\ s^* \in \mathcal{K},\ x^* \circ s^* = 0$$

which implies that x^* and s^* operator commute (cf. [5, Proposition 6]). Therefore, they also admit a simultaneous spectral decomposition. The next lemma relates the simultaneous spectral decomposition (5.10) with the simultaneous spectral decomposition of x^* and s^*.

Lemma 5.9. *Consider the simultaneous decomposition of $x(\mu)$ and $s(\mu)$ given by (5.10). Then, there exist a simultaneous spectral decomposition of x^* and s^*, namely,*

$$x^* = \sum_{i=1}^{r} \gamma_i^* c_i^*, \qquad s^* = \sum_{i=1}^{r} \delta_i^* c_i^* \tag{5.11}$$

and a sequence $\{\mu_k\} \subset \mathrm{int}\mathbb{R}_+$ satisfying $\mu_k \downarrow 0$ when $k \to \infty$ such that for every $i \in \{1, \ldots, r\}$ one has

$$c_i(\mu_k) \to c_i^*, \qquad \gamma_i(\mu_k) \to \gamma_i^*, \qquad \delta_i(\mu_k) \to \delta_i^*, \tag{5.12}$$

when $k \to \infty$.

Proof. The continuity of the Jordan product and the fact that the set of all primitive idempotents in a simple EJA is compact (cf. [8]) lead us to conclude that the set of Jordan frames is compact. Since the central path is bounded, we also have that the eigenvalues of $x(\mu)$ and $s(\mu)$ are bounded. Hence, the set

$$\{c_1(\mu), \ldots, c_r(\mu), \gamma_1(\mu), \ldots, \gamma_r(\mu), \delta_1(\mu), \ldots, \delta_r(\mu) : \mu > 0\} \tag{5.13}$$

is bounded. Therefore, there exist sequences given by (5.12) whose limit point satisfy the simultaneous spectral decomposition given in (5.11). □

Let $u \in \mathbb{V}$ be such that it admits the spectral decomposition $u = \sum_{i=1}^{r} \lambda_i d_i$. We introduce the following index notation

$$I := \{1, \ldots, r\}, \; I_+(u) := \{i \in I : \lambda_i > 0\}, \; I_0(u) := \{i \in I : \lambda_i = 0\}.$$

The Jordan frame in (5.11) can be partitioned as

$$\{c_i^* : i \in I_+(x^*)\} \cup \{c_i^* : i \in I_+(s^*)\} \cup \{c_i^* : i \in I_0(x^*) \cap I_0(s^*)\}.$$

We define

$$c_B^* := \sum_{i \in I_+(x^*)} c_i^*, \quad c_T^* := \sum_{i \in I_0(x^*) \cap I_0(s^*)} c_i^*, \quad c_N^* := c_B^* + c_T^*. \tag{5.14}$$

The Peirce decompositions of $\mathbb{V}$ with respect to the idempotents c_B^* and c_N^* are given, respectively, by

$$\mathbb{V} = \mathbb{V}(c_B^*, 1) \oplus \mathbb{V}(c_B^*, 1/2) \oplus \mathbb{V}(c_B^*, 0), \tag{5.15}$$

$$\mathbb{V} = \mathbb{V}(c_N^*, 1) \oplus \mathbb{V}(c_N^*, 1/2) \oplus \mathbb{V}(c_N^*, 0). \tag{5.16}$$

We use the words *primal* and *dual* Peirce decomposition of $\mathbb{V}$ to refer the decomposition (5.15) and (5.16), respectively.

Example 5.10. We illustrate the case $\mathbb{V} = \mathbb{S}^n$ (semidefinite programming). Let (X^*, y^*, S^*) be the limit point of the central path. We denote

$$|B| := \mathrm{rank}(X^*) \quad \text{and} \quad |N| := \mathrm{rank}(S^*).$$

Note that in general $|B| + |N| \leq n$. Without loss of generality (applying an orthonormal transformation of problem data, if necessary) we can assume that

$$X^* = \begin{bmatrix} \Lambda_B^* & 0 \\ 0 & 0 \end{bmatrix}, \quad S^* = \begin{bmatrix} 0 & 0 \\ 0 & \Lambda_N^* \end{bmatrix},$$

where Λ_B^* and Λ_N^* are positive definite diagonal matrices of order $|B|$ and $|N|$, respectively. Hence, we can choose as common Jordan frame the set $\{e_1e_1^T, \ldots, e_ne_n^T\}$ where $\{e_1, \ldots, e_n\}$ is the canonical basis of $\mathbb{R}^n$. Therefore, the idempotents defined in (5.14) becomes

$$C_B^* = \begin{bmatrix} I_{|B|} & 0 \\ 0 & 0 \end{bmatrix}, \quad C_N^* = \begin{bmatrix} I_{n-|N|} & 0 \\ 0 & 0 \end{bmatrix},$$

where $I_{|B|}$ and $I_{n-|N|}$ are the identity matrices of order $|B|$ and $n - |N|$, respectively. Furthermore, we have that

$$\mathbb{V}(C_B^*, 1) = \left\{ \begin{bmatrix} U & 0 \\ 0 & 0 \end{bmatrix} : U \in \mathbb{S}_+^{|B|} \right\}, \quad \mathbb{V}(C_N^*, 0) = \left\{ \begin{bmatrix} 0 & 0 \\ 0 & V \end{bmatrix} : V \in \mathbb{S}_+^{|N|} \right\}.$$

The next proposition shows that the partitions (5.15) and (5.16) are optimal in the sense that $\mathbb{V}(c_B^*, 1)$ and $\mathbb{V}(c_N^*, 0)$ are subalgebras of $\mathbb{V}$ with the smallest rank containing the primal and dual optimal sets, respectively.

Proposition 5.11. *We have that*

$$\mathcal{S}_P \subseteq \mathbb{V}(c_B^*, 1), \quad \mathcal{S}_D \subseteq \mathbb{R}^m \times \mathbb{V}(c_N^*, 0).$$

Proof. Let us consider the simultaneous spectral decompositions of x^* and s^* and the sequence $\{\mu_k\}$ given by Lemma 5.9. Let $(x, y, s) \in \mathcal{S}_P \times \mathcal{S}_D$. We get from the orthogonality relation (5.8) that

$$\langle x(\mu_k) - x, s(\mu_k) - s \rangle = 0. \tag{5.17}$$

By using the fact that $x \circ s = 0$ and $x(\mu_k) \circ s(\mu_k) = (\mu_k/\rho)e$ one has

$$\langle x, s(\mu_k) \rangle + \langle x(\mu_k), s \rangle = r\mu_k.$$

Moreover, since $x(\mu_k)^{-1} = \frac{\rho}{\mu_k} s(\mu_k)$ and $s(\mu_k)^{-1} = \frac{\rho}{\mu_k} x(\mu_k)$ the las equality becomes

$$\langle x, x(\mu_k)^{-1} \rangle + \langle s, s(\mu_k)^{-1} \rangle = \rho r. \tag{5.18}$$

Observe that the two terms of the left hand of the last equation are nonnegative. This implies that

$$\langle x, x(\mu_k)^{-1} \rangle \leq \rho r, \qquad \langle s, s(\mu_k)^{-1} \rangle \leq \rho r. \tag{5.19}$$

By using the spectral decomposition of $x(\mu_k)$ one obtains that $x(\mu_k)^{-1} = \sum_{i=1}^r (1/\gamma_i(\mu_k)) c_i(\mu_k)$. Hence, the first inequality of (5.19) becomes

$$\sum_{i=1}^r \frac{1}{\gamma_i(\mu_k)} \langle x, c_i(\mu_k) \rangle \leq \rho r. \tag{5.20}$$

Since each term of the left hand side of (5.20) is nonnegative, we deduce that

$$0 \leq \langle x, c_i(\mu_k)\rangle \leq \rho r \gamma_i(\mu_k), \ \forall i \in I.$$

Letting $k \to +\infty$, this implies that

$$0 \leq \langle x, c_i^*\rangle \leq \rho r \gamma_i^*, \ \forall i \in I.$$

Therefore, one obtains

$$\langle x, c_i^*\rangle = 0, \ \forall i \in I_0(x^*). \tag{5.21}$$

From [5, Proposition 6], and the fact that $x, c_i^* \in \mathcal{K}$ it follows that (5.21) is equivalent to

$$x \circ c_i^* = 0, \ \forall i \in I_0(x^*).$$

Finally, since $c_B^* + \sum_{i \in I_0(x^*)} c_i^* = e$, one gets

$$x \circ c_B^* = x \circ \left(e - \sum\nolimits_{i \in I_0(x^*)} c_i^*\right) = x.$$

That is $x \in \mathbb{V}(c_B^*, 1)$. We have thus proved that $\mathcal{S}_P \subseteq \mathbb{V}(c_B^*, 1)$. The proof of $\mathcal{S}_D \subseteq \mathbb{R}^m \times \mathbb{V}(c_N^*, 0)$ is analogous. □

Definition 5.12. *We say that* $(\bar{x}, \bar{y}, \bar{s}) \in \mathcal{S}_P \times \mathcal{S}_D$ *is a* maximally complementary solution *if it maximizes* $\operatorname{rank}(x) + \operatorname{rank}(s)$ *over* $\mathcal{S}_P \times \mathcal{S}_D$. *Recall that* $\operatorname{rank}(x)$ *stands for the number of nonzeros eigenvalues of* x.

Corollary 5.13. *The limit point of the central path is a maximally complementary solution.*

Proof. The result follows from Proposition 5.11 since it implies the following stronger relation

$$\operatorname{rank}(x) \leq \operatorname{rank}(x^*), \forall x \in \mathcal{S}_P, \quad \operatorname{rank}(s) \leq \operatorname{rank}(s^*), \forall (y, s) \in \mathcal{S}_D$$

□

The next proposition characterizes of the primal and dual optimal sets in SCLP.

Proposition 5.14. *The primal and dual optimal sets in SCLP can be characterized, respectively, as*

$$\mathcal{S}_P = \mathcal{F}_P \cap \mathbb{V}(c_B^*, 1), \tag{5.22}$$
$$\mathcal{S}_D = \mathcal{F}_D \cap (\mathbb{R}^m \times \mathbb{V}(c_N^*, 0)). \tag{5.23}$$

Proof. The inclusion $\subseteq$ was proved in the Proposition 5.11. Let us prove the reverse inclusion $\supseteq$ for (5.22). Equality (5.23) follows analogously. Let $x \in \mathcal{F}_P \cap \mathbb{V}(c_B^*, 1)$ and $(y, s) \in \mathcal{S}_D$ we claim $\langle x, s\rangle = 0$. Indeed,

$$\langle x, s\rangle = \langle x \circ c_B^*, s\rangle = \langle x, s \circ c_B^*\rangle. \tag{5.24}$$

Since $s \in \mathbb{V}(c_N^*, 0)$ we have that

$$s \circ c_B^* = s \circ (c_N^* - c_T^*) = -s \circ c_T^*.$$

This together with (5.24) implies

$$\langle x, s\rangle = -\langle x, s \circ c_T^*\rangle = -\langle x \circ c_T^*, s\rangle.$$

Finally, since $x \in \mathbb{V}(c_B^*, 1)$, $c_T^* \in \mathbb{V}(c_B^*, 0)$ and $\mathbb{V}(c_B^*, 1) \circ \mathbb{V}(c_B^*, 0) = \{0\}$ (cf. [2, Proposition IV.1.1]), we conclude that $x \circ c_T^* = 0$, which leads to $\langle x, s\rangle = 0$. We have thus proved that (x, y, s) satisfies the optimality conditions (5.1)-(5.3), i.e., $x \in \mathcal{S}_P$. □

For $\tau \in \{B, N\}$ and $\alpha \in \{0, 1/2, 1\}$, we introduce the following notations

$$\begin{cases} P_{\tau,\alpha} \text{ denotes the projection onto } \mathbb{V}(c_\tau^*, \alpha), \\ \mathcal{K}_{\tau,\alpha} := \{u^2 : u \in \mathbb{V}(c_\tau^*, \alpha)\}, \\ \mathrm{int}_{\tau,\alpha} C \text{ denotes the interior of the set } C \subseteq V(c_\tau^*, \alpha) \\ \text{with respect to the topology in } V(c_\tau^*, \alpha). \end{cases}$$

Note that, there exists a collection $\{a_1, \ldots, a_m\} \subset \mathbb{V}$ such that the linear map $\mathcal{A} : \mathbb{V} \to \mathbb{R}^m$ can be expressed as

$$\mathcal{A}x = (\langle a_1, x\rangle, \ldots \langle a_m, x\rangle)^T. \tag{5.25}$$

By using the projections onto the primal and dual Peirce decomposition, Proposition 5.14 can be written in the following way:

Proposition 5.15. *The primal and dual optimal sets in SCLP can be characterized, respectively, as*

$$\mathcal{S}_P = \{x \in \mathcal{K}_{B,1} : \langle P_{B,1}(a_i), x\rangle = b_i,\ i = 1, \ldots, m\},$$
$$\mathcal{S}_D = \left\{(y, s) \in \mathbb{R}^m \times \mathcal{K}_{N,0} : \sum_{i=1}^m y_i P_{N,0}(a_i) + s = P_{N,0}(c),\ \sum_{i=1}^m y_i P_{N,\alpha}(a_i) = P_{N,\alpha}(c),\ \alpha = 1/2, 1\right\}.$$

Proof. Taking the representation (5.25) for the linear map $\mathcal{A}$ and noting that $\mathbb{V}(c_\tau^*, \alpha) \cap \mathcal{K} = \mathcal{K}_{\tau,\alpha}$ ($\tau \in \{B, N\}$, $\alpha \in \{0, 1\}$), the result of Proposition 5.14 can be expressed as

$$\mathcal{S}_P = \{x \in \mathcal{K}_{B,1} : \langle a_i, x\rangle = b_i\,(i = 1, \ldots, m)\}, \tag{5.26}$$
$$\mathcal{S}_D = \{(y, s) \in \mathbb{R}^m \times \mathcal{K}_{N,0} : \sum_{i=1}^m y_i a_i + s = c\}. \tag{5.27}$$

We show the characterization of $\mathcal{S}_P$. For every $i \in \{1, \ldots, m\}$, the decomposition of a_i with respect to the primal Peirce decomposition (5.16) is given by

$$a_i = P_{B,1}(a_i) + P_{B,1/2}(a_i) + P_{B,0}(a_i).$$

The fact that $x \in \mathbb{V}(c_B^*, 1)$ implies that

$$\langle P_{B,1/2}(a_i), x\rangle = \langle P_{B,0}(a_i), x\rangle = 0.$$

Hence, $\langle a_i, x\rangle = \langle P_{B,1}(a_i), x\rangle$ and the relation (5.26) becomes

$$\mathcal{S}_P = \{x \in \mathcal{K}_{B,1} : \langle P_{B,1}(a_i), x\rangle = b_i,\ i = 1, \ldots, m\}\,.$$

Analogously, in order to show the characterization of $\mathcal{S}_D$, we decompose a_i and c with respect to the dual Peirce decomposition (5.16):

$$\begin{aligned} a_i &= P_{N,1}(a_i) + P_{N,1/2}(a_i) + P_{N,0}(a_i), \\ c &= P_{N,1}(c) + P_{N,1/2}(c) + P_{N,0}(c). \end{aligned}$$

Taking into account the above decompositions and the fact that $s \in \mathbb{V}(c_N^*, 0)$ imply that the condition $\sum_{i=1}^m y_i a_i + s - c = 0$ is equivalent to

$$\left(\sum_{i=1}^m y_i P_{N,1}(a_i) - P_{N,1}(c)\right) + \left(\sum_{i=1}^m y_i P_{N,1/2}(a_i) - P_{N,1/2}(c)\right) + \left(\sum_{i=1}^m y_i P_{N,0}(a_i) + s - P_{N,0}(c)\right) = 0.$$

This is equivalent to

$$\sum_{i=1}^m y_i P_{N,\alpha}(a_i) = P_{N,\alpha}(c),\ \alpha = 1/2, 1, \quad \text{and} \quad \sum_{i=1}^m y_i P_{N,0}(a_i) + s = P_{N,0}(c).$$

Hence, the relation (5.27) becomes

$$\mathcal{S}_D = \left\{(y, s) \in \mathbb{R}^m \times \mathcal{K}_{N,0} : \sum_{i=1}^m y_i P_{N,0}(a_i) + s = P_{N,0}(c),\ \sum_{i=1}^m y_i P_{N,\alpha}(a_i) = P_{N,\alpha}(c),\ \alpha = 1/2, 1\right\}$$

□

The above result allows us to describe the relative interior of the primal and dual optimal sets in SCLP.

Corollary 5.16. *The relative interior of the primal and dual optimal sets in SCLP are give, respectively, by*

$$\begin{aligned} \mathrm{ri}(\mathcal{S}_P) &= \{x \in \mathcal{S}_P : x \in \mathrm{int}_{B,1}\mathcal{K}_{B,1}\}, \\ \mathrm{ri}(\mathcal{S}_D) &= \{(y, s) \in \mathcal{S}_P : s \in \mathrm{int}_{N,0}\mathcal{K}_{N,0}\}. \end{aligned}$$

In particular, the limit point of the central path in SCLP lies in the relative interior of $\mathcal{S}_P \times \mathcal{S}_D$.

5.5 Conclusions and further work

The results presented in this paper extend known results on the central path in semidefinite programming. We think that this unifying approach enlights known results for particular EJA. For instance, the Peirce decomposition helps us to clarify the block notation that are present in some results of SDP (cf. Example 5.10).

A full characterization of the limit points of the central path is still an open question. We believe that, as occurs in SDP, under the strictly complementarity assumption the limit point should coincide with the analytic center of the primal-dual optimal set.

Our future works will intend to characterize this limit point by following the approach presented in [7, Theorem 3.2]. It was shown therein that the limit point of the central path in SDP is the analytic center of some subset of the primal-dual optimal set. We also expect to study the central path in symmetric cone convex programming associated with a larger class of penalty-barrier functions as it was carried out in [10] for the SDP case.

Bibliography

[1] M. Baes. Spectral functions and smoothing techniques on Jordan algebras. PhD. Thesis, Université catholique de Louvain, 2006.

[2] J. Faraut and A. Korányi. *Analysis on Symmetric Cones*. Clarendon Press, Oxford, 1994.

[3] L. Faybusovich. Euclidean Jordan algebras and interior-point algorithms. *Positivity 1* 4 (1997), 331–357.

[4] L. Faybusovich. Linear systems in Jordan algebras and primal-dual interior point algorithms. *J. Comput. Appl. Math.* 86 (1997), 149–175.

[5] M.S. Gowda, R. Sznajder, and J. Tao. Some P-properties for linear transformations on Euclidean Jordan algebras. *Linear Algebra Appl.*, 393 (2004), 203–232.

[6] M. Halická, E. De Klerk and C. Roos. On the convergence of the central path in semidefinite optimization. *SIAM J. Optim.* 12 (2002), 1090–1099.

[7] M. Halická, E. De Klerk and C. Roos. Limiting behavior of the central path in semidefinite optimization. *Optim. Methods Softw.* 20 (2005), 99–113.

[8] U. Hirzebruch. Über Jordan-Algebren und kompakte Riemannsche symmetrische Räume vom Rang 1. *Math. Z.* 90 (1965), 339–354.

[9] E. de Klerk, C. Roos and T. Terlaky. Infeasible-start semidefinite programming algorithms via self-dual embeddings. *Fields Inst. Commun.,* 18 (1998), 215–236.

[10] J. López and H. Ramírez. On the central paths and Cauchy trajectories in semidefinite programming. *Kybernetika* 46 (2010), 524–535.

[11] Y. E. Nesterov and M. J. Todd. Self-scaled barriers and interior-point methods for convex programming. *Math. of Oper. Res.* 22 (1997), 1–42.

[12] Y. E. Nesterov and M. J. Todd. Primal-dual interior-point methods for self-scaled cones. *SIAM J. Optim.* 8 (1998), 324–364.

[13] S. H. Schmieta and F. Alizadeh. Extension of primal-dual interior point algorithms to symmetric cones. *Math. Program.* 96 (2003), 409–438.

[14] G. Q. Wang and Y. Q. Bai. A new full Nesterov-Todd step primal-dual path-following interior-point algorithm for symmetric optimization. *J. Optim. Theory Appl.* 154 (2012), 966–985.

[15] J. Zhang and K. Zhang. Polynomial complexity of an interior point algorithm with a second order corrector step for symmetric cone programming. *Math. Methods Oper. Res.* 73 (2011), 75–90.